A.A. LOGUNOV

GRAVITATION AND ELEMENTARY PARTICLE PHYSICS

A.A. LOGUNOV

GRAVITATION AND ELEMENTARY PARTICLE PHYSICS

PREFACE

It is a quest to unify four, seemingly different, fundamental interactions: strong, electromagnetic, weak, and gravitational, that characterize modern field theory. Yet until recently the prospects for such a unification seemed all too vague. The last few years have however witnessed some solid progress towards the unified description of the weak and electromagnetic interactions based on the gauge approach and these studies were given a new powerful push.

It is a discussion of some of the tasks that need solving before this goal can be approached that is now offered to our readers.

Having characterized the current status of field theory, it is essential to note above all that the theoretical field concepts used to describe gravitational interaction are radically different from those used to describe weak, strong, and electromagnetic interactions.

It is well known that the most popular comprehensive theory of gravitation, i.e. Einstein's general theory of relativity (GTR), actually moves away from the conventional understanding of "field" by identifying the field variable with the metric tensor of Riemannian space-time, thus making the theory completely geometric. As a result, the GTR has proved to be devoid, in principle, of conservation laws for matter and gravitational field when taken in conjunction. Many attempts, even going back to the work of Einstein himself, to introduce into the GTR a substitute for the energy-momentum conservation law of matter and gravitational field taken in conjunction have all, when closely scrutinized, proved unsuccessful. The GTR is in fact constructed at the expense of the conservation laws and so the gravitational field in Einstein's theory is not in the spirit of Faraday-Maxwell. Such a radical difference between the gravitational field and other fields would be justified had there been telling reasons and primarily had there been experimental data indicating non-conservation of energy-momentum in any physical process. However, there are no such data and this signifies that, in the case of gravitational interaction too, we have no reason to reject a fundamental physical law, i.e. that of energy-momentum conservation. Hence the task of constructing a theory that would, on one hand, allow a gravitational field to be treated as an energy-momentum carrier analogous to any other physical field, and that would, on the other hand, also conserve Einstein's grand idea of a space-time geometry

closely associated with matter. Solving this problem would permit the abyss that has, until now, separated the theory of gravitation from those of the other interactions to be bridged, and thus undoubtedly help the unification of these theories.

This is one of the possibilities offered by the field approach to the construction of a new theory of gravitation and which is analyzed in the first paper "The theory of space-time and gravitation" by V.I. Denisov and A.A. Logunov. This approach is based on two fundamental propositions. Firstly, it is required that the conservation laws hold for matter and gravitational field taken in conjunction, this being achieved by choosing a pseudo-Euclidean geometry as the inherent one for the gravitational field. Using this approach, a gravitational field is characterized by its own energy-momentum tensor which contributes to the system's complete energy-momentum tensor.

Secondly, there is a geometrization principle (an identity principle) which states that the equations of motion of matter under the action of a gravitational field in the pseudo-Euclidean space-time have a metric tensor γ_{ni} that can be equated with identical equations of motion of matter in an effective Riemannian space-time having a metric tensor g_{ni} which depends on the gravitational field and the metric tensor γ_{ni}. This identity principle reflects the universal nature of the interaction of gravitational field with other physical fields and it follows from the results of gravitational experiments although a definite choice of the Lagrangian density of interactions between all the fields must be made to describe it. According to this principle, the covariant energy-momentum conservation law of matter and gravitational field in the pseudo-Euclidean space-time can be represented as a covariant energy-momentum conservation equation for the matter alone in the effective Riemannian space-time. This proves that an effective Riemannian geometry emerges within the scope of the field approach because, figuratively speaking, the gravitational field possesses energy-momentum.

One of the most natural implementations of the field approach is a theory of gravitation using a symmetric tensor of second rank to describe the gravitational field. Since this field involves four irreducible representations having the spins 2, 1, 0, and 0′, the formalism of projection operators is used to eliminate the surplus spin states, thus ensuring, in turn, the gauge invariance of the theory.

This field theory of gravitation allows all the available gravitational experiments in the solar system to be described and makes it possible to construct a non-stationary model of the homogeneous Universe that can explain the observed consequences of the cosmological expansion of the Universe. In addition, it satisfies the principle of correspondence to the Newtonian gravitation theory, and finally, it predicts a set of fundamental corollaries. It should be noted, however, that the theory suggested is only one of the many

theoretical schemes conceivable. The experimental evidences of the various interactions are rather distinctly specified in terms of their own characteristic regularities. Hence, one of the most peculiar features of strong interactions is the multiple production of particles at high energies, which means that multiple production processes play a principal role in the study of the nature of strong interactions.

These processes have been discovered in cosmic rays and for a long time semiphenomenological concepts based on applications of the laws of thermo- and hydrodynamics had to be employed to explain their mechanism. The pioneering theoretical studies of the characteristics of multiple processes such as the number density of particles, the energy, etc., are associated with the names of W. Heisenberg, E. Fermi, L.D. Landau, I.Ya. Pomeranchuk *et al.* and date back to the late forties and early fifties. At that time the study of multiple production processes of particles within the framework of quantum field theory looked very promising since the amplitudes of the corresponding transitions were extremely complicated and the experimental research encountered fundamental difficulties. The problem is that kinematics of events with a large number of particles is extremely complex and there were no particular hopes of studying it in any detail using standard characteristics (chosen reaction channels and their differential cross sections), nor were there any particular ways to study them. At the then comparatively low energies of the colliding particles, the number of reaction products was small; however, new and powerful particle accelerators were commissioned in the sixties. The fraction of multiple processes has unremittingly risen as the energy was increased and there was an urgent need to find a simpler and more convenient method to describe multiple production processes.

Such a method was found in 1967 in the work by A.A. Logunov, M.A. Mestverishvili, and Nguyen Van Hiew. A production cross section of a single secondary particle is introduced as the main object of the theory instead of their transition amplitudes, whilst integrating over the variables related to the remaining secondaries. As a result, we obtain a quantity (now known as the inclusive cross section) which can be investigated both theoretically and experimentally in a much simpler way than the ordinary cross sections and which, at the same time, contains sufficient information about the dynamics of interacting particles and their structure. In a similar fashion production cross sections of two, three, and more chosen particles can be introduced and, following Feynman's proposal in 1969, every process contributing to the production of the given number of the chosen (detectable) particles is called inclusive.

The inclusive approach in high energy physics has proved to be most suitable for the description of the multiple production processes that have no dependence on the type of interacting particles and it has in many ways promoted the development of this branch of

physics. In fact, the introduction of inclusive cross sections has immediately allowed the powerful apparatus of axiomatic quantum field theory, which was developed earlier but had only been applied before to the processes of elastic scattering and charge exchange, to be applied to multiple production processes as well. A number of substantial restrictions that are imposed on the high energy behavior of inclusive cross sections have been obtained using these general principles of quantum field theory. Even in the first experiments on measuring inclusive cross sections a previously unknown regularity that is characteristic of inclusive processes as such was discovered, i.e. the scale invariance of inclusive cross sections. In the course of successive experiments an appreciable number of these scaling laws was discovered for the multiple production processes of various origins. The advance of the inclusive approach has also stimulated the development of many phenomenological models which enable experimental results to be both described and predicted.

The need for a theoretical substantiation of the scaling in deep inelastic processes by the field theory has provided a new and powerful impetus for a thorough investigation of the non-Abelian gauge theories that has resulted in the significant advance in this area of research. Some extremely important problems have now also appeared on the agenda, such as the relation between the inclusive cross sections and three-particle scattering amplitudes and hence the limitations on the behavior of these cross sections need to be established. A survey of the results obtained in the last few years by studying inclusive processes, both starting from general principles and also within the scope of various models, is presented in the second paper, "Inclusive processes and the dynamics of strong interactions", by A.A. Logunov, M.A. Mestverishvili, and V.A. Petrov.

The third paper, by N.N. Bogolyubov, M.A. Matveev, and A.N. Tavkhelidze, "Colored quarks", contains a review of a number of the most important advances in elementary particle physics, nuclear physics, and high-energy physics that have been inferred from the concept that colored quarks are fundamental constituents of matter. The notion of color, i.e. a new quantum number, was introduced in 1965 by N.N. Bogolyubov, B.V. Struminskii, and A.N. Tavkhelidze in the USSR and independently, by Y. Nambu and M.Y. Han in the USA, in connection with the problem of quark statistics. It is now basic to hadron spectroscopy and quantum chromodynamics as well as the various versions of the unified gauge theories of strong, weak, and electromagnetic interactions.

At the outset the problem of the dynamic description of hadrons as composite quark systems is discussed as is a construction of the form factors and amplitudes of various processes involving hadrons. The main aim of building dynamic quark models is to explain why quarks have not, despite many attempts, been discovered in a free state.

Explaining the absence of free quarks is called quark confinement or non-emission and remains to date one of the most important tasks of elementary particle physics and quantum field theory. The dynamic model initiated at JINR (Dubna) in 1964 was based on the idea that heavy quarks were bound within hadrons by immense forces, that, on one hand, dictates a large mass defect of quarks in hadrons and, on the other, prevents their emission. These ideas have boosted the development of the modern quark models of elementary particles, the quark bag model and the parton model being the most popular.

As will be shown in the paper, the dynamic composite model enable both the observed static features of elementary particles (magnetic moments, axial-vector constants of weak transitions, etc.) and the hadron form factors to be described systematically. Notice in particular that the enhancement of the magnetic moments of a heavy quark bound in a hadron has for the first time been satisfactorily explained within the framework of this model, and has enabled the absolute values of the proton and neutron magnetic moments (in nuclear magnetons) to be evaluated:

$$\mu_p \simeq 3, \quad \mu_n \simeq -2.$$

In addition, the model can describe the mass splittings that occur within meson and baryon multiplets and permits the renormalized axial constant of the nucleon weak interaction, and its relativistic corrections, to be determined allowing for the internal quark motion in hadrons. A quark model of the electromagnetic and weak meson decays, which was developed from the dynamic approach, was vital for the elaboration of elementary particle theory. Thus the weak lepton decays of the pseudoscalar π- and K-mesons and the electromagnetic decays of the vector ρ^0-, ω^0-, and φ^0-meson resonances into electron-positron pairs can be described as the annihilation of quarks and antiquarks bound in these mesons. Notice that the relevant decay widths are governed by the magnitudes of the functions of bound quark-antiquark pairs for matching coordinates:

$$\Gamma(\pi \to \mu\tilde{\nu}) = \frac{G^2\cos^2\theta}{2\pi^2} m_\mu^2 \left(1 - \frac{m_\mu^2}{m_\pi^2}\right)^2 |\psi_\pi(0)|^2,$$

$$\Gamma(V^0 \to e^+e^-) = \frac{16\pi\alpha^2}{3m_V^2} Q_V^2 |\psi_V(0)|^2,$$

where

$$Q_\rho = \frac{1}{\sqrt{2}}, \quad Q_\omega = \frac{1}{3\sqrt{2}}, \quad Q_{\varphi^0} = -\frac{1}{3}.$$

These formulas and the idea of the quark annihilation mechanism itself are at the base of the current theoretical analysis of the various decay modes of the members of a new family of heavy particles, namely, ψ/J- and γ-mesons. Y. Nambu made an important con-

tribution to dynamic hadron theory when, for the first time in 1965, he introduced vector fields, i.e. the carriers of the color interaction which were prototypes of quantum-chromodynamical gluon fields. It should be stressed that quantum chromodynamics (QCD), whose rapid progress over the last decade has been observed by us, emerged as a result of generalizing the color $SU^c(3)^{(g)}$ symmetry as a local gauge transformation.

It is scarcely possible to enumerate all the advances of QCD whose development was significant for the progress of strong interaction theory. As will be shown in detail in this paper, QCD has meant that a wide range of phenomena having approximate scaling invariance (the automodel behavior) can be treated consistently and a quark counting method for processes with high transferred momenta can be well-grounded. It is hoped that quark confinement can be theoretically explained by the non-Abelian nature of the QCD gauge symmetry. In the last few years colored quarks and fundamental QCD forces have begun to find their way into the theory of nuclear phenomena. Let us emphasize that the asymptotic behavior pattern of the deuteron electromagnetic form factor is a direct indication of a quark structure in the nucleus, as it matches quark counting predictions quite well. It is well known that a quark counting formula, that was established in 1973, governs the energy dependence of the differential cross sections for large angle scattering and the hadron form factors at high energies, $E = \sqrt{S}$, and high momentum transfers, $Q = \sqrt{-t}$:

$$\frac{d\sigma}{dt} = (ab \to cd) \left(\frac{1}{S}\right)^{n_a+n_b+n_c+n_d-2} f(\theta),$$

$$F_a(t) \sim \left(\frac{1}{t}\right)^{n_a-1},$$

where $n_{i=a,b,c,d}$ are the numbers of elementary constituents (quarks and antiquarks) in the reacting hadrons and θ is the scattering angle in the center-of-inertia frame. The quark counting formula describes the numerous experimental data on elementary particle scattering surprisingly well and makes it possible for explicit information about the number of hadrons' elementary constituents to be deduced from the experiments. This information, when applied to an analysis of recent experimental data on electron-deuteron scattering, indicates the presence of a hard six-quark deuteron structure.

This paper discusses the function of quark degrees of freedom when describing nuclear phenomena, especially those occurring at high energies and momentum transfers and indicates, in particular, the possible excitation of "hidden color" in nuclear matter and a number of other corollaries. It seems certain, the notion that colored quarks and gluons are fundamental constituents of matter will radically change our ideas about the world of atomic nuclei and a new light

will be shed both on the properties of nuclear matter and the nature of nuclear matter and nuclear forces.

The analysis of Greenberg's hypothesis about a quark's para-fermi statistics performed by the authors of this paper is interesting. Whilst the hypothesis does enable a problem, in which there is a virtual violation of the Pauli principle for a baryon consisting of three quarks, to be solved, para-fermi statistics is too narrow and does not permit the gauge SU^c (3) symmetry that forms the QCD basis to be introduced.

A maximum gauge symmetry compatible with a parastatistics of rank 3 is the SO (3) group. It has three gluons and a particle spectrum containing diquarks and other exotic hadrons.

The paper is concluded by a discussion about whether color symmetry is an exact or approximate law of nature. This principal problem of elementary particle theory, as yet unsolved, is closely related to the question about quark charges.

Even the first works on the three-triplet model indicated that in the case of colored quarks integral values could be chosen for the electric and baryonic charges. Introducing integral quark charges which are dependent on the quark's color state results in an obvious breakdown of color symmetry, at least for electromagnetic interactions of particles.

This paper considers the unified gauge models of the strong and electromagnetic interactions which spontaneously break color symmetries and integral charge quarks, and discusses the corollaries of these models that are being observed experimentally.

These profound and principal problems, dealt with in this paper, obviously cannot be solved theoretically, and experiment must have the last word.

It should be stressed that the hypothesis about the integrity of quark charges resulted, in particular, in the concept of unstable quarks and was the starting point for a number of unified gauge models of elementary particles tolerating nucleon decay and other processes in which the baryon number is not conserved. An experimental verification of the predictions of similar theories should take place soon.

This paper, by N.N. Bogolyubov, V.A. Matveev, and A.N. Tavkhelidze, should help scientists to be aware of the profound influence the idea that colored quarks are fundamental constituents of matter has exerted on the development of the physics of elementary particles and nuclear and high energy physics, and assists the evaluation of the qualitative changes that have been observed in those areas of research in the last two decades.

3 March 1982

A.A. *Logunov*

1 The Theory of Space-Time and Gravitation

V. I. Denisov and A. A. Logunov

1.1 INTRODUCTION

Einstein's general theory of relativity (GTR) is one of the fundamental physical theories at the present time and has a deep rooted concept of a relation between matter and space embedded in it. The theory explained and predicted a number of gravitational effects, a genuine triumph.

However, there are a number of difficult problems with the GTR that have remained unsolved, one of which is the basic one of energy-momentum of a gravitational field. A general study of this problem [1-9] has led us to conclude that it is impossible, in principle, to solve it within the framework of the GTR since the gravitational field in Einstein's theory is not really a field in the spirit of Faraday-Maxwell, i.e. it is not characterized by an energy-momentum tensor density. This can be verified easily by comparing the physical properties of gravitational and other fields.

All the physical theories that describe the different forms of matter include energy-momentum tensor density as one of the most important features of a field. This tensor density is commonly obtained by varying the density of the field Lagrangian, L, with respect to the components of a metric space-time tensor g_{ni}[1] thus:

$$T^{ni} = -2\frac{\delta L}{\delta g_{ni}} = \sqrt{-g}\,\mathsf{T}^{ni} \tag{1.1}$$

where T^{ni} is the field energy-momentum tensor. This feature shows a field exists since a nonzero value of the density of the energy-momentum tensor is a necessary and sufficient condition for a physical field in this region. The energy-momentum of any physical field contributes to the complete energy-momentum tensor of a system and does not become zero outside the field source. This makes it possible to consider energy transfer by waves in the Faraday-Maxwell spirit, i.e. to study the field intensity pattern in space, to determine the energy fluxes through a surface, to compute the changes in the energy-momentum value during radiation and absorption, and to perform other energy based computations.

Henceforth the Latin indices run over the values 0, 1, 2, 3, and the Greek ones, 1, 2, 3. The metric signature is chosen in the form (+, —, —, —).

The gravitational field in the GTR does not have the features typical of other physical fields, for it does not have the feature discussed above.

Indeed, the density of Lagrangian in Einstein's theory is composed of two parts, the density of the gravitational field Lagrangian, $L_g = L_g(g_{ni})$, which depends only on the metric tensor g_{ni}, and the density of the matter Lagrangian, $L_M = L_M(g_{ni}, \varphi_A)$, which depends on the metric tensor g_{ni} and the remaining fields of the matter φ_A. Thus, the quantities g_{ni} in Einstein's GTR have two meanings, being both the variables of the field and the metric space-time tensor.

As a result of this physical and geometrical duality, an expression for the density of a complete, symmetric energy-momentum tensor (a variation of the Lagrangian density with respect to metric-tensor components) coincides with field equations (a variation of the Lagrangian density with respect to gravitational field components). This implies that the density of the complete, symmetric energy-momentum tensor of a system is strictly equal to zero:

$$T^{ni} + t^{ni} = 0, \tag{A}$$

where $T^{ni} = -2\frac{\delta L_M}{\delta g_{ni}}$ is the density of the symmetric energy-momentum tensor of matter (here by matter we assume all other fields, too, except gravitational),

$$t^{ni} = -2\frac{\delta L_g}{\delta g_{ni}} = -\frac{c^4\sqrt{-g}}{8\pi G}\left[R^{ni} - \frac{1}{2}g^{ni}R\right]. \tag{1.2}$$

It also follows from expression (A) that all the components of the density of the symmetric energy-momentum tensor t^{ni} of the gravitational field vanish everywhere outside matter.

These results clearly demonstrate that the gravitational field in Einstein's GTR does not exhibit the properties typical of other physical fields, since it does not have that basic physical feature, an energy-momentum tensor outside the source.

A curvature tensor R^i_{nlm} is a physical characteristic of a gravitational field in Einsten's theory. A clear explanation is given by Synge [10, p. VIII]: "... If we accept the idea that space-time is a Riemannian four-space (and if we are relativists we must), then surely our first task is to get the feel of it just as early navigators had to get the feel of a spherical ocean. And the first thing we have to get the feel of is the Riemann tensor, for it is the gravitational field—if it vanishes, and only then, there is no field. Yet, strangely enough, this most important element has been pushed into the background." And further he wrote: "... In Einstein's theory, either there is a gravitational field or there is none, according as the Riemann tensor does not or does vanish. It is an absolute property; it has nothing to do with any observer's world-line..."

Unfortunately, some of the theorists who specialize in the GTR have still not understood this fundamental point. This lack of understanding might explain, for example, statements by a number of authors [11, 16, 17, 74, 127] who claim that, given a transformation into an appropriate coordinate system, a gravitational field within a small space-time region could be considered as absent in Einstein's theory. Yet a physical feature of gravitational fields is in fact that they can change the energy-momentum of matter, i.e. it demonstrates the force action of a gravitational field on matter described by equation [11] thus:

$$\frac{\delta^2 n^i}{\delta s^2} + R^i_{mkl}\, u^m u^l n^k = 0, \tag{1.3}$$

where $u^i = \frac{dx^i}{ds}$ is the velocity four-vector, and n^i is the infinitesimal displacement vector of geodesics. Yet a description using curvature waves yields no information concerning a wave transferred energy flux.

Thus, Einstein's GTR combines matter and a gravitational field, the first being characterized, as in all other physical theories, by the energy-momentum tensor (a tensor of rank two), and the second, by the curvature tensor (a tensor of rank four). From the difference in the dimensions of physical parameters of the gravitational field and matter in Einstein's theory it follows immediately that there are, in principle, no GTR conservation laws relating matter and gravitational field. This fundamental fact established first by us [6] suggests that Einstein's theory has been constructed at the expense of the laws of conservation, rejecting them when matter and gravitational field are taken together.

H.A. Lorentz and T. Levi-Civita have proposed the quantities in (1.2) to be considered as the density components of the energy-momentum tensor of gravitational field, and expression (A) as a specific conservation law for the density of complete energy-momentum tensor. The conservation law (A) is peculiar because it is a local conservation law enabling the change in the energy-momentum tensor of gravitational field at some point to be found from the change in the energy-momentum tensor of matter at that point:

$$\frac{\partial}{\partial t} T^{0i} = -\frac{\partial}{\partial t} t^{0i}. \tag{1.4}$$

However, in Einstein's theory, the tensor t^{ni} is only characteristic of geometry inside matter, hence a change in the GTR energy-momentum of the matter is only directly related to a change in the scalar curvature R and the tensor R^{ni} of rank two within the region occupied by the matter. The curvature waves described by the tensor R^i_{nlm} of rank four are not directly related to the change in the energy-momentum of matter, but are related implicitly via the metric

tensor g_{ni}. Because of this, there are no conservation laws for the GTR curvature waves, relating the change in the energy-momentum tensor of matter (a tensor of rank two) to the change in the curvature tensor (a tensor of rank four). So, the Riemannian space-time in the GTR is, on one hand, a source of energy-momentum, since the curvature tensor affects the particle motion according to equation (1.3) while the energy of the Riemannian space-time, on the other hand, is produced without complying with the conservation laws for the energy-momentum of matter and gravitational field when taken in conjunction.

Einstein was not satisfied with the introduction of the conservation law based on expression (A). He wrote [12a]: "...Indeed, it is impossible to advance any *logical* objections against such a *denomination.* However, I think that it is impossible to draw the same conclusions from equation (A) that we are accustomed to do from the conservation laws. This is due to the fact that, according to (A), the components of a *total-energy tensor* vanish everywhere" (the authors' italics). Further, Einstein emphasized that according to (A), a material system could be completely dissolved without leaving any traces behind, since its energy (A) was zero.

Einstein correctly points out that it is impossible to draw the sort of conclusions from equation (A) as can be done from the conservation laws, yet it is not just a case here of a *name* but rather of the essence of the GTR.

However, Einstein thought that, in the GTR, the gravitational field with matter should have a certain conservation law [12b]: "...surely, one has to require both matter and gravitational field to satisfy together the energy-momentum conservation laws".

He tried to discover conservation laws for matter and gravitational field that would be analogous to the conservation laws in classical mechanics or electromagnetic field theory. As is generally known, this forced him to introduce a noncovariant quantity (the energy-momentum pseudotensor) into the covariant theory and the formal analogue of the conservation laws of classical mechanics and electrodynamics could only be obtained at this price. In order to obtain this type of "conservation laws" [11], one usually acts in the following manner.

Given Einstein's equations in the form

$$-\frac{c^4}{8\pi G} g\left[R^{ik}-\frac{1}{2} g^{ik}R\right]=-gT^{ik}, \tag{1.5}$$

the left-hand side might be represented as the sum of two noncovariant quantities:

$$-\frac{c^4}{8\pi G} g\left[R^{ik}-\frac{1}{2} g^{ik}R\right]=\frac{\partial}{\partial x^l} h^{ikl}+g\tau^{ik}, \tag{1.6}$$

where τ^{ik} is the energy-momentum pseudotensor of gravitational field, and $h^{ikl} = -h^{ilk}$ is the spin pseudotensor.

Using identity (1.6), Einstein's equations (1.5) may be written in the equivalent form:

$$-g\,[T^{ik}+\tau^{ik}]=\frac{\partial}{\partial x^l}\,h^{ikl}. \tag{1.7}$$

Due to the obvious equality

$$\frac{\partial^2}{\partial x^k\,\partial x^l}\,h^{ikl}=0,$$

a differential conservation law can be deduced from Einstein's equations in the following form:

$$\frac{\partial}{\partial x^k}\,[-g\,(T^{ik}+\tau^{ik})]=0. \tag{1.8}$$

Integrating this relation over some volume and assuming the absence of material fluxes through the surface enclosing this volume, one generally obtains from expression (1.8) the integral "energy-momentum conservation laws", too [12]:

$$\frac{d}{dt}\int(-g)\,[T^{0i}+\tau^{0i}]\,dV=-\oint(-g)\,\tau^{\alpha i}\,dS_\alpha. \tag{1.9}$$

Einstein [12c] suggested that the right-hand side of this relation, when $i=0$, "... is surely the energy lost by the material system". In the absence of "energy-momentum fluxes" of a gravitational field through the surface enclosing the volume of integration, one obtains the conservation law for the "energy-momentum" of a system from expression (1.9):

$$P^i=\frac{1}{c}\int(-g)\,[T^{0i}+\tau^{0i}]\,dV=\text{const}. \tag{1.10}$$

Using Einstein's equations (1.7), relation (1.10) can be rewritten as

$$P^i=\frac{1}{c}\oint h^{i0\alpha}\,dS_\alpha=\text{const}. \tag{1.11}$$

According to Einstein [12d], the four quantities P^i are the energy ($i=0$) and the momenta ($i=1,2,3$) of a physical system. It is generally stated thereby [11] that "... the quantities P^i, a four-momentum of field and matter, have a definite meaning, i.e. they appear to be independent of the particular choice of coordinate system to the extent necessary from a physical standpoint". However, as we shall demonstrate below, this statement is wrong.

Analogous results are also obtained when Einstein's equations are written in terms of mixed components:

$$\sqrt{-g}\,[T^i_k+\tau^i_k]=\partial_n\sigma^{ni}_k. \tag{1.12}$$

The choice of the energy-momentum pseudotensors for gravitational field depends to a great extent on the personal inclination of the authors and was, as a rule, made on the basis of the secondary properties. Thus, for example, h^{ikl} is chosen in the form

$$h^{ikl}=\frac{c^4}{16\pi G}\frac{\partial}{\partial x^m}[-g(g^{ik}g^{ml}-g^{il}g^{km})], \tag{1.13}$$

a symmetrical Landau-Lifshitz pseudotensor is obtained which contains only the first derivatives of the metric tensor.

Choosing

$$\sigma_k^{ni}=\frac{g_{km}c^4}{16\pi G\sqrt{-g}}\frac{\partial}{\partial x^l}[-g(g^{im}g^{nl}-g^{nm}g^{il})], \tag{1.14}$$

we arrive at Einstein's pseudotensor which coincides with a canonical energy-momentum (pseudo)tensor obtained from the noncovariant density of the gravitational field Lagrangian:

$$L_g=\sqrt{-g}\,[\Gamma^n_{mk}\Gamma^l_{nl}-\Gamma^n_{lm}\Gamma^l_{nk}]\,g^{mk}.$$

With

$$\sigma_k^{ni}=\frac{c^4\sqrt{-g}}{16\pi G}g^{im}g^{nl}[\partial_l g_{km}-\partial_m g_{kl}], \tag{1.15}$$

we get the Lorentz pseudotensor which coincides with a canonical energy-momentum (pseudo)tensor obtained with a noncovariant method of infinitesimal displacements from the covariant density of the gravitational field Lag angian $L_g = \sqrt{-g}R$. It is important to note here that, however diversified the energy-momentum pseudotensors might be, in terms of their properties they all possess the same feature, i.e. any energy-momentum pseudotensor can be made to vanish at any point in space.

This fact can be interpreted as a reflection of the equivalence principle, however, any statement concerning the equivalence of a gravitational field and a field of inertial forces is an error. These two fields are substantially different since, given an inertial force field, the curvature tensor always vanishes but it has a nonzero value for a gravitational field. As a consequence, an inertial force field and a gravitational field are not equivalent for all physical processes that are strongly dependent on the curvature tensor. Hence the equivalence principle does not bear any direct relation to the GTR even though it played an heuristic role during its construction by Einstein.

The fact that a pseudotensor vanishes at any given point in space is a consequence of a nontensor law that transforms its components from one coordinate system to another. Thus, for example, all the energy-momentum pseudotensors containing derivatives of the metric tensor, in a Riemannian space-time, of an order not greater

than one, vanish on being transformed to a local geodetic coordinate system where all the components of the connections Γ^i_{nm} are zero. Thus the "energy-momentum" of gravitational field determined from the energy-momentum pseudotensors can be made to vanish locally.

Yet a gravitational field, described by a curvature tensor, cannot be made to vanish simply by a transformation to another arbitrary, admissible[2] coordinate system, and, consequently, due to the action of curvature waves on physical processes, it is impossible to state a gravitational field is absent in any coordinate system. Because of this, as has been indicated in [1-9], the energy-momentum pseudotensors are not energy-momentum characteristics of a gravitational field and are not indicative of its existence either locally or globally, this resulting in physically senseless definitions "energy-momentum" or "energy fluxes" of the system in the GTR.

This general conclusion has been illustrated in [8] by citing, as an example, the GTR's definition of "inertial mass" and the "energy" of a static, spherically symmetric system. Based on definition (1.10) of the "energy-momentum" of a system composed of matter and gravitational field, the notion of an inertial mass m of a system is introduced into the GTR as follows:

$$m=\frac{1}{c}P^0=\frac{1}{c^2}\int(-g)\,[T^{00}+\tau^{00}]\,dV=\frac{1}{c^2}\oint h^{00\alpha}\,dS_\alpha. \quad (1.16)$$

To calculate a system's inertial mass, the Schwarzschild solution is usually used and hence the metric of the Riemannian space-time can be written in terms of isotropic Cartesian coordinates in the form

$$\begin{aligned} g_{\alpha\beta} &= -\delta_{\alpha\beta}\left[1+\frac{r_g}{4r}\right]^4, \\ g_{00} &= \left[1-\frac{r_g}{4r}\right]^2\left[1+\frac{r_g}{4r}\right]^{-2}, \end{aligned} \quad (1.17)$$

where $r=\sqrt{x^2+y^2+z^2}$, $r_g=2GM/c^2$, and M is the gravitational or, as it is sometimes called, heavy mass. These coordinates are asymptotically Galilean, since at $r\to\infty$ the following approximations are true:

$$g_{00}=1+O\left(\frac{1}{r}\right);\quad g_{\alpha\beta}=-\delta_{\alpha\beta}\left[1+O\left(\frac{1}{r}\right)\right]. \quad (1.18)$$

---

[2] By admissible transformations we understand coordinate transformations between reference systems, which can be performed with the help of real physical bodies and processes. Mathematically, this condition is equivalent [13] to a requirement that a quadratic form (with the coefficients $g_{\alpha\beta}$) exists in these reference systems which is negative by definition, and a component g_{00} of the metric tensor is positive: $g_{00}>0$, $g_{\alpha\beta}\,dx^\alpha\,dx^\beta<0$.

Using the covariant components of metric (1.17) we obtain from (1.11)

$$P^0 = \frac{c^3 r_g}{2G} = Mc. \tag{1.19}$$

It is this identity between the "inertial mass" and the heavy one that gave rise to statements about their equality in the GTR [11, p. 424]: "... $P^\alpha = 0$, $P^0 = Mc$ is a result which has naturally been expected. It illustrates the equality of the so-called "heavy" and "inertial" masses. (One calls a mass "heavy" when it determines a gravitational field produced by a body; it is the mass entering the metric tensor in the gravitational field or, in particular, Newton's law. The "inertial" mass determines a relationship between momentum and energy of the body and, in particular, the body rest energy is equal to this mass times c^2.)"

Yet a similar statement made by Einstein [12e] and other authors is wrong. As is shown in our paper [8], the "energy" of system (1.10) and, as a consequence, its "inertial" mass (1.16) too, have no physical meaning since their value even depends on the choice of the three-dimensional coordinate system.

It is quite evident that it is an elementary requirement of the definition of inertial mass, which any physical theory has to satisfy, that the value of material mass is independent of the choice of three-dimensional coordinate system. However, definition (1.16) of the inertial mass in the GTR does not satisfy this requirement. Suppose, for example, we transform the three-dimensional Cartesian coordinates x^α_{old} into the new coordinates x^α_{new}, the two being connected via the expression

$$x^\alpha_{\text{old}} = x^\alpha_{\text{new}}\,[1 + f\,(r_{\text{new}})], \tag{1.20}$$

where $r_{\text{new}} = \sqrt{x^2_{\text{new}} + y^2_{\text{new}} + z^2_{\text{new}}}$, and $f\,(r_{\text{new}})$ is an arbitrary non-singular function satisfying the conditions

$$f\,(r_{\text{new}}) \geqslant 0,\ \lim_{r_{\text{new}}\to\infty} f\,(r_{\text{new}}) = 0,$$
$$\lim_{r_{\text{new}}\to\infty} r_{\text{new}} \frac{\partial}{\partial r_{\text{new}}} f\,(r_{\text{new}}) = 0. \tag{1.21}$$

It can easily be shown that transformation (1.20) corresponds to a change in the arithmetization of the points of the three-dimensional space along the radius:

$$r_{\text{old}} = r_{\text{new}}\ [1 + f\,(r_{\text{new}})].$$

In order that transformation (1.20) has an inverse transformation and is one-to-one, it is necessary and sufficient for it to fulfil the condition

$$\frac{\partial r_{\text{old}}}{\partial r_{\text{new}}} = 1 + f + r_{\text{new}} f' > 0,$$

where $f' = \frac{\partial}{\partial r_{\text{new}}} f(r_{\text{new}})$.

The transformation Jacobian will then be nonzero:

$$J = \det \left\| \frac{\partial x_{\text{old}}}{\partial x_{\text{new}}} \right\| = (1+f)^2 \frac{\partial r_{\text{old}}}{\partial r_{\text{new}}} \neq 0.$$

In particular, all the above requirements are satisfied by the function

$$f(r_{\text{new}}) = \alpha^2 \sqrt{\frac{8GM}{c^2 r_{\text{new}}}} [1 - \exp(-\varepsilon^2 r_{\text{new}})], \tag{1.22}$$

where α and ε are arbitrary numbers with nonzero values. Since here

$$\frac{\partial r_{\text{old}}}{\partial r_{\text{new}}} = 1 + \alpha^2 \sqrt{\frac{8GM}{c^2 r_{\text{new}}}} \left[\frac{1}{2} + \left(\varepsilon^2 r_{\text{new}} - \frac{1}{2} \right) \exp(-\varepsilon^2 r_{\text{new}}) \right],$$

$f(r_{\text{new}})$ is a monotonic function of r_{new}. It can easily be seen that $f(r_{\text{new}})$ is non-negative and non-singular over the entire space. In this case the transformation Jacobian is strictly greater than unity:

$$J = (1+f)^2 \frac{\partial r_{\text{old}}}{\partial r_{\text{new}}} > 1.$$

Therefore, transformation (1.20) with the function $f(r_{\text{new}})$ determined by expression (1.22) has an inverse transformation and is one-to-one.

Now let us compute the value of the "inertial mass" (1.16) in terms of the new coordinates x^{α}_{new}. Taking advantage of the transformation law of the metric tensor

$$g^{\text{new}}_{ik}(x_{\text{new}}) = \frac{\partial x^l_{\text{old}}}{\partial x^i_{\text{new}}} \frac{\partial x^m_{\text{old}}}{\partial x^k_{\text{new}}} g^{\text{old}}_{lm}(x_{\text{old}}(x_{\text{new}})), \tag{1.23}$$

we shall find the components of the Schwarzschild metric in terms of the new coordinates. As a result, we obtain

$$\begin{aligned} g_{00} &= \left[1 - \frac{r_g}{4r_{\text{new}}(1+f)}\right]^2 \left[1 + \frac{r_g}{4r_{\text{new}}(1+r)}\right]^{-2}, \\ g_{\alpha\beta} &= \left[1 + \frac{r_g}{4r_{\text{new}}(1+f)}\right]^4 \Big\{ -\delta_{\alpha\beta}(1+f)^2 \\ &\quad - x^{\text{new}}_{\alpha} x^{\text{new}}_{\beta} \left[(f')^2 + \frac{2}{r_{\text{new}}} f'(1+f) \right] \Big\}. \end{aligned} \tag{1.24}$$

The determinant of the metric tensor (1.24) is equal to

$$g=-g_{00}\left[1+\frac{r_g}{4r_{\text{new}}(1+f)}\right]^{12}(1+f)^4[(1+f)^2+r_{\text{new}}^2(f')^2$$
$$+2r_{\text{new}}f'(1+f)]. \qquad (1.25)$$

It should be noted that metric (1.24) is asymptotically Galilean, since

$$\lim_{r_{\text{new}}\to\infty} g_{00}=1,\ \lim_{r_{\text{new}}\to\infty} g_{\alpha\beta}=-\delta_{\alpha\beta}.$$

In the particular case of the function $f(r_{\text{new}})$ given by relation (1.22) and $r_{\text{new}}\to\infty$, the metric of the Riemannian space-time will have the following asymptotics:

$$g_{00}=1+O\left(\frac{1}{r_{\text{new}}}\right),\quad g_{\alpha\beta}=-\delta_{\alpha\beta}+O\left(\frac{1}{\sqrt{r_{\text{new}}}}\right). \qquad (1.26)$$

Expressions (1.18) and (1.26) show that even the rate at which the three-dimensional part of the metric tensor of the Riemannian space-time tends to a Galilean value depends on the choice of the space coordinates, i.e. it depends on the arithmetization mode of the points in space rather than on some physical condition. Therefore, any requirement imposed on the asymptotic behavior of the three-dimensional part of a metric tensor is not physical but is dictated only by one or another method of the arithmetizing the points in space. Yet a theory should always guarantee the ability to choose any admissible coordinate system. Thus even a restriction, in the form of some chosen method of the arithmetization of the spatial points, is a senseless requirement.

Substituting the contravariant components of the metric (1.24) into expression (1.16), we obtain

$$m=\frac{c^2}{2G}\lim_{r_{\text{new}}\to\infty}\{r_g+r_{\text{new}}^3(f')^2\}. \qquad (1.27)$$

Thus, the value of the "inertial mass" is substantially dependent on the rate at which f' tends to zero as $r_{\text{new}}\to\infty$. In particular, choosing the function $f(r_{\text{new}})$ in the form of (1.22) and solving expression (1.27) for the "inertial mass" we get:

$$m = M(1+\alpha^4). \qquad (1.28)$$

Since the arbitrariness in the value of α is associated in the GTR with the arbitrariness in the choice of the spatial coordinates, we can obtain an arbitrarily prescribed value $m\geqslant M$ for the "inertial mass" (1.16) of a system composed of matter and gravitational field even though the gravitational mass M of this system and, consequently, all three GTR effects too, remain unaltered thereby. Note also that under more complicated transformations of the spatial coordinates, which leave the metric asymptotically Galilean, the

"inertial mass" (1.16) of the system can take any permissible value, both positive and negative.

Thus we see that the GTR value of the "inertial mass", being primarily introduced by Einstein and subsequently adopted by many authors [11, 14-18], depends on the choice of the three-dimensional coordinate system and, as a consequence, it has no physical meaning. Therefore, statements concerning the identity of the inertial and heavy masses in Einstein's theory also have no physical meaning. The identity only holds for a narrow class of the three-dimensional coordinate systems and, as the "inertial" (1.16) and gravitational masses are transformed in different ways, their identity does not hold when transformed to another three-dimensional coordinate system. Besides, the GTR definition of the inertial mass (1.16) does not satisfy the principle of correspondence to Newton's theory. Indeed, since the inertial mass m in Einstein's theory depends on the choice of the three-dimensional coordinate system, its expression, in the general case of an arbitrary three-dimensional coordinate system, cannot be transformed into a matching expression in Newton's theory where the inertial mass is independent of the choice of the spatial coordinates. Thus, the classical Newtonian limit is absent in the GTR and, as such, the GTR does not satisfy the correspondence principle.

A similar situation is observed, too, when defining the "energy fluxes" of gravitational radiation. The "intensity of gravitational radiation" is defined in the GTR using the components $\tau^{0\alpha}$ of the energy-momentum pseudotensor:

$$\frac{dI}{d\Omega} = -\lim_{r\to\infty} c\,(-g)\,r^2\,\tau^{0\alpha}n_\alpha. \tag{1.29}$$

Due to a substantial nonlinearity of Einstein's equations, an investigation of the wave solutions is generally limited to a first approximation in small-wave perturbations. In this case the following expression, in terms of the asymptotically Cartesian coordinates, is obtained for the "intensity of gravitational radiation":

$$\frac{dI}{d\Omega} = \frac{G}{36\pi c^5}\left\{\frac{1}{4}(\dddot{D}_{\alpha\beta}\,n^\alpha n^\beta)^2 + \frac{1}{2}\dddot{D}_{\alpha\beta}\,\dddot{D}^{\alpha\beta} + \dddot{D}_{\alpha\beta}\,\dddot{D}^{\alpha\gamma}\,n^\beta n_\gamma\right\}, \tag{1.30}$$

where

$$D^{\alpha\beta} = \int dV\,(3x^\alpha x^\beta - \gamma^{\alpha\beta}x_\varepsilon x^\varepsilon)\,T^{00}.$$

Integrating expression (1.30) over a sphere, we obtain the well known Einstein quadrupole formula which is usually used to prove the "reality" of the existence of an "energy flux" of gravitational waves emitted by a radiating island system:

$$I = -\frac{dE}{dt} = \frac{G}{45c^5}\dddot{D}_{\alpha\beta}\,\dddot{D}^{\alpha\beta}. \tag{1.31}$$

So, expressions (1.30) and (1.31) seem to support a conclusion made by Einstein [12f]: "... It can be seen from the formula... that the radiation intensity cannot be negative in any direction, the more so for the total intensity of radiation..."

Yet this conclusion too, as has been shown in [9], is erroneous. The value of the "gravitational radiation intensity" (1.30) in the GTR as well as the "total intensity" according to Einstein are substantially dependent on the choice of coordinates. Therefore, by an appropriate and admissible transformation of coordinates (see Sec. 1.10), which leaves the Riemannian space-time metric asymptotically Galilean at infinity, these quantities can vanish or even become negative in a spatial region between two spheres with radii $r_1 = ct - u_1$ and $r_2 = ct - u_2$. Hence it appears that the "gravitational radiation intensity" (1.30) and the "total intensity" (1.31) are not energy-momentum characteristics of a gravitational field in the GTR, since the radiation, being an objective physical reality, cannot be destroyed by any admissible transformation of coordinates. So, for example, in electrodynamics, as can be readily shown, the energy flux of electromagnetic radiation cannot vanish due to a transformation of coordinates. If the energy flux of electromagnetic waves through a surface is nonzero in one reference frame, then after a transformation to another admissible reference frame it cannot vanish or, moreover, change sign.

As a consequence, formula (1.31) for the energy lost by a source due to gravitational radiation is not, in principle, contained in the GTR, since there is no inherent way to make any energy calculations.

The "energy flux" of gravitational radiation, defined using energy-momentum pseudotensors which can vanish in the lowest order of perturbation theory given an appropriate choice of coordinates, is a reflection of a general statement that a coordinate system can be chosen for any energy-momentum pseudotensor such that the "energy flux" of gravitational radiation is always strictly zero. The coordinate system that has this property can be obtained by reducing the components $g^{0\alpha}$ of the metric tensor to the form that satisfies the condition

$$\partial_n \sigma_0^{n\alpha} \equiv 0$$

for energy-momentum pseudotensors with the mixed components. As a result and due to expressions (1.12) and (1.9), the "energy flux" of gravitational waves in these coordinate systems will be absent:

$$\frac{d}{dt} \int dV \, (-g)^{1/2} \, [T_0^0 + \tau_0^0] \equiv 0.$$

In a similar manner appropriate coordinate systems can be established for other types of energy-momentum pseudotensors as well.

Yet the given coordinate systems preserve all the solutions of Einstein's equations with the nonzero curvature tensor and, as a

consequence, the curvature waves capable of imparting energy-momentum to physical bodies do exist in these coordinate systems. This statement may be illustrated quite simply with a Lorentz pseudotensor and as is evident from (1.15) and (1.12) we may nullify the "total energy-momentum density" of a gravitational field by a transformation to a synchronous reference system ($g_{00} = 1$, $g_{0\alpha} = 0$):

$$\sqrt{-g}\,[T^i_0 + \tau^i_0] \equiv 0.$$

Hence the "energy-momentum" of gravitational field should be absent outside matter, too. Yet the curvature waves are solutions of Einstein's equations and exist in the synchronous reference system. They affect physical processes by changing their energy-momentum.

Recently a set of papers [19-21] have appeared claiming to have obtained an expression for the mass of a system composed of matter and a gravitational field in the GTR using a Hamiltonian formalism and to have its positive definiteness. A hasty conclusion has even been drawn [126] on the basis of these statements, that the problem of the energy-momentum of gravitational field in Einstein's theory was solved. However, these statements only testify to the misunderstanding of the essence of the problem the authors of [19-21, 126] have. In fact, it can easily be shown that all these studies [19-21] really rely on a requirement that the asymptotic behavior of the metric tensor in Riemannian space-time conforms to same law at space infinity [126], thus:

$$g_{ik} = \delta_{ik} + O\left(\frac{1}{r}\right), \quad \partial_n g_{ik} = O\left(\frac{1}{r^2}\right). \tag{1.32}$$

It is this requirement which enables them to obtain an expression for the GTR "mass" of a system and to prove its positive definiteness. Yet this requirement is not a physical one. As can readily be seen with the Schwarzschild solution which exhibits an asymptotic behavior (1.18) in terms of isotropic Cartesian coordinates (1.17) that satisfies (1.32).

If the points of three-dimensional space are arithmetized by some other method (and the arithmetization of the three-dimensional space is always arbitrary so all theories should allow for the arbitrariness of the choice of arithmetization), then we generally obtain another law governing the asymptotic behavior of the spatial part of the metric tensor of the Riemannian space-time.

In particular case considered above, after transformation (1.20) is performed, the components of the metric tensor were shown to have the following asymptotics:

$$g_{00} = 1 + O\left(\frac{1}{r}\right) \text{ and } g_{\alpha\beta} = -\delta_{\alpha\beta} + O\left(\frac{1}{\sqrt{r}}\right).$$

Hence it follows that the asymptotic behavior of the three-dimensional part of the metric tensor in the Riemannian space-time is determined by the space points arithmetization method rather than by any physical requirement. The change in the asymptotics of the three-dimensional part of the metric tensor that occurs when the arithmetization of the space points changes totally invalidates the immense efforts expended by the authors of [19-21] to prove the mathematical statements. Mathematical proofs are undoubtedly the most important constituents of theoretical physics, but proof is only meaningful when the problem is formulated in a physically correct form. Otherwise proofs, however elegant, have no value at all in physics. As academician A.N. Krylov was wont to say [125], "... Mathematics, like millstone, grinds everything poured under it and, just as you won't get wheat flour by grinding Deadly Nightshade, you won't get the truth from the false premises even if you cover the page with formulae..." This applies with a vengeance to the above work. On one hand, [19-21, 126] are incorrect because of the physical formulation of the problem itself, while, on the other hand, the expression used for the "mass" of the system (see, e.g. [126]):

$$m = \oint \partial_i q^{ik} dS_k = \oint \frac{\partial}{\partial x^m} [-g(g^{00}g^{im} - g^{0i}g^{0m})]\, dS_i,$$

is also explicitly dependent on the method of arithmetizing the three-dimensional spatial points, in other words, it is not scalar with respect to the choice of the three-dimensional coordinate system. Naturally, this is physically senseless, since the value of the system's "mass" may be made equal to any number desired by an appropriate choice of the three-dimensional coordinate system.

It should also be noted that an approach being the Hamiltonian is philosophically close to pseudotensor formalism which is obvious from the above formula for a system's "mass". We have indicated the reasons for the errors in [19-21, 126] and we might add that the authors of [19-21, 126] did not appreciate the fundamental finding that there is in principle no conservation law in the GTR when the energy-momentum of matter and gravitational field are taken together. It is impossible, therefore, to introduce a notion of a system energy-momentum into the GTR either.

At the end of Secs. 1.10 and 1.12 we shall return again to this problem and explain why the GTR gives, in some coordinate systems, physically admissible formulas which, nevertheless, are not contained in the theory.

The other approach to the GTR energy-momentum problem, and which has mainly been applied to approximate calculations, is based on the production of the so-called motion integrals from the equations of motion of matter and based on the covariant conservation equation for the matter energy-momentum tensor. Using this approach, a nonconservation of matter energy, if discovered at some

point in an approximation calculation, is generally explained [13, 127-130] as the radiation of gravitational waves by matter, and thus their "energy" can be determined as the "force" of the gravitational radiation friction can.

Results obtained in this way are contradictory. For example, in [22-24] it was concluded, that the gravitational wave energy has the negative sign, since the system energy had increased as the gravitational waves were emitted from this system. At the same time, similar works [25-27] show a decrease in the system's energy when gravitational waves are emitted and, as a consequence, these waves should transfer a positive energy.

Yet from the covariant equation for the conservation of the matter energy-momentum tensor one can only obtain a trivial conservation law of type (1.4) in a rigorous way. In fact, due to the conservation equation for the energy-momentum tensor density of matter:

$$\nabla_n T^{ni} = \partial_n T^{ni} + \Gamma^i_{nm} T^{nm} = 0, \tag{1.33}$$

we have

$$\nabla_n [(-g)^\alpha T^n_i] = (-g)^\alpha g_{ik} \nabla_n T^{nk} = 0.$$

Hence it follows that

$$\partial_n \{(-g)^\alpha T^n_i\} = (-g)^\alpha [\Gamma^m_{ni} T^n_m + 2a\Gamma^m_{nm} T^n_i].$$

Making use of Einstein's equations

$$T^n_i = -t^n_i,$$

where

$$t^n_i = -\frac{c^4 \sqrt{-g}}{8\pi G}\left[R^n_i - \frac{1}{2}\delta^n_i R\right],$$

we obtain

$$\partial_n \{(-g)^\alpha T^n_i\} = (-g)^\alpha [-\Gamma^m_{ni} t^n_m - 2a\Gamma^m_{nm} t^n_i]. \tag{1.34}$$

It follows from

$$\nabla_m R^m_n = \frac{1}{2}\frac{\partial}{\partial x^n} R$$

that

$$(-g)^\alpha [\Gamma^m_{ni} t^n_m + 2a\Gamma^m_{nm} t^n_i] = \partial_n [(-g)^\alpha t^n_i].$$

Substituting this expression into the right-hand side of equation (1.34), we get a conservation law

$$\partial_n \{(-g)^\alpha [T^n_i + t^n_i]\} = 0. \tag{1.35}$$

Analogously, we can obtain a conservation law in the form

$$\partial_n \{(-g)^\alpha [T^{ni} + t^{ni}]\} = 0. \tag{1.36}$$

From expression (1.36) and according to equations (1.2) and (1.6), two differential equations also appear:

$$\partial_n \partial_m h^{inm} = 0,$$
$$\partial_n \{(-g)^{\alpha} [T^{ni} + \tau^{ni}]\} = 0,$$

which both reflect only a local fulfilment of Einstein's equations rather than being some sort of conservation law.

Thus, the covariant conservation equation (1.33) and Einstein's equations lead us to relations (1.35) and (1.36), the latter being trivial as they are satisfied by the field equations and (1.4) is just a special case of these relations.

As has been shown in [6], the ambiguous results obtained from approximations of the "energy" and "force" of the gravitational radiation friction are simply a consequence of the arbitrary transfer of the tensor terms t^{0i} in expression (1.4) from the right-hand side to the left-hand side which is concluded by declaring that the right-hand side of the final expression is the energy flux of gravitational waves. It is quite evident that this procedure is absolutely meaningless and produces different results depending on what is transferred to the left, viz. a positive or negative quantity.

To sum up:

(a) The GTR has not, and cannot have, energy-momentum conservation laws when a gravitational field and matter are taken in conjunction.

(b) The inertial mass defined in the GTR has no physical meaning.

(c) Einstein's quadrupole formula for gravitational radiation is not a corollary of the GTR.

(d) It does not follow from the GTR, in principle, that a double star system loses its energy by gravitational radiation.

(e) The GTR does not have the classical Newtonian limit and, hence, it does not satisfy a fundamental physical principle, that is the correspondence principle.

Thus the gravitational field in the GTR differs entirely from other physical fields and is not a field in the Faraday-Maxwell spirit. All these taken together indicate that the GTR is not a satisfactory physical theory. It should be noted that the GTR is only one of the possible realizations of Einstein's grand idea about the Riemannian space-time geometry. So, whilst discussing the inadequacy of Einsteinian theory, bear in mind it is only the inadequacy of that particular realization. Inasmuch as the theories of other physical fields, a unified conservation law of energy-momentum exists for the different forms of matter, and since there is at present no experimental evidence of its violation (moreover, the history of physics has always illustrated its tenacity and truth), there is no reason to reject this law. Therefore, we shall assume that a conservation law relating the energy-momentum of different forms of matter must be the basis of

any physical theory. Only experimental evidence could make us reject such an assumption and the law must be valid for all fields, including the gravitational one. For these reasons, constructing a theory of gravitation which satisfies all the requirements for the physical theory is vital at present.

What ways are open? What can be retained from Einstein's tremendous work and what has to be rejected to satisfy the fundamental physical laws of a new gravitation theory, viz. the energy-momentum conservation law when matter and a gravitational field are taken in conjunction, and the correspondence principle?

To answer these questions, we shall consider the ideas that form the basis of Einstein's GTR.

The most profound of these is, in our opinion, the idea of a Riemannian geometry of space-time whose metric tensor g_{ni} is determined by matter. Another hypothesis supporting the whole construction of the GTR is that the gravity and space-time metric are united. This is accomplished by describing gravity using the metric tensor g_{ni}.

As has clearly been shown by Hilbert, these two hypotheses lead, in the simplest case, to Einstein's famous GTR equations. Since the GTR departed from the conventional concepts of a gravitational field in the spirit of Faraday-Maxwell, then, in order to construct a new theory of gravitation that is similar to the theories of other physical fields and has the conventional property that a gravitational field is an energy-momentum carrier, we have to retain and enrich Einstein's first idea and reject his second. This is the way we chose.

Our work [28-33] was dedicated to this problem and contained a formulation of field theory of gravitation. During last years, however, our views have evolved somewhat and so our previous publications [1-7, 28-31] have become, in a sense, steps to the concepts we now hold and which are offered in the present article.

1.2 SPACE-TIME GEOMETRY AND CONSERVATION LAWS

When a physical theory has a tensor quantity as a field variable, the form of the differential field equations should be independent of the choice of coordinates used. This may be achieved in two ways, either by allowing only the covariant derivatives in the space-time metric inherent in the process to be present in the field equations, or by composing the tensor quantity from the field functions and their partial derivatives. In the latter case the field equations will be substantially nonlinear.

When Einstein constructed the GTR, he used the second way connecting the Riemannian space-time metric tensor g_{ni} and matter

using the nonlinear equations (A). As a result, the idea that matter could influence space-time metric was produced.

As we have seen, however, such an approach does not permit the GTR's gravitational field to be considered as a physical one which possesses energy-momentum. Besides, the natural geometry for gravitational field in the GTR was a Riemannian space-time geometry. This was not the result of experimental evidence but rather a hypothesis about the character of a gravitational field's self-action. Yet a gravitational field self-action need not reduce to a mere change in the geometry albeit nonlinear. As a consequence, a question arises as to which inherent geometry for gravitational field will enable the field to possess an energy-momentum density.

Any physical field has an appropriate inherent geometry such that, in the absence of an interaction with other fields, a free wavefront of this physical field moves along the geodesic of the inherent space-time.

The propagation equation for a wavefront of a massless field (the characteristic equation) [11],

$$g^{ni}\frac{\partial\psi}{\partial x^n}\frac{\partial\psi}{\partial x^i}=0, \tag{1.37}$$

and the equation of motion of free material particles (the Hamilton-Jacobi equation),

$$g^{ni}\frac{\partial\psi}{\partial x^n}\frac{\partial\psi}{\partial x^i}=1, \tag{1.38}$$

are both defined by the metric tensor of the geometries inherent in these processes.

The problem of choosing which inherent geometry to use boils down to finding out which effective metric tensor controls the convolution of the higher-order derivatives in the Lagrangian density. There is a conceivable situation, already mentioned by Lobachevsky [34], in which all the different physical phenomena are described by different inherent geometries.

Equations (1.37) and (1.38) suggest that the inherent geometry of a physical theory could allow for an experimental definition based on data about motion of test particles and fields. The study of motion of test particles with mass in massless fields permits the metric tensor to be obtained in the inherent space-time to within some constant factor [35]. Thus, the study of the motion of different forms of matter make it possible to verify the character of the world's space-time geometry experimentally. For this reason, our understanding of space-time has advanced as our knowledge of nature has developed.

So, Newtonian mechanics (mechanical phenomena) coupled with the Galilean principle of relativity (as we know it now) has once again established that space is Euclidean and time is absolute, i.e. the same in all coordinate systems. Yet the connection between the

Galilean relativity principle and the geometry of Newtonian mechanics was not established and it was therefore considered to be an independent principle only pertaining to inertial coordinate systems. Initially, this principle was only associated with mechanical phenomena and, subsequently, was extended by Poincaré [36] to all physical phenomena when formulated in the following way: "... the laws of physical phenomena must be the same for a "fixed" observer as for an observer who has a uniform motion of translation relative to him: so that we have not, and cannot possibly have, any means of discerning whether we are, or are not, carried along in such a motion".

It should be noted that although this principle appears natural, its true nature was unclear.

Later on, Faraday-Maxwell electrodynamics (electromagnetic phenomena) in combination with the principle of relativity brought about the discovery of a pseudo-Euclidean geometry of the world space-time. This is largely due to Minkowski, who wrote in his book "Space and Time" [37]: "... The views of space and time which I wish to lay before you have sprung from the soil of experimental physics, and therein lies their strength. They are radical. Henceforth space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality," and, later on, he noted that "... it is only in four dimensions that the relations here taken under considerations reveal their inner being in full simplicity, and that on a three-dimensional space forced upon us a priori they cast only a very complicated projection...".

Minkowski was the first to discover that the space-time, in which all physical processes occur, is unified and has a pseudo-Euclidean geometry. Subsequent study of strong, electromagnetic, and weak interactions has demonstrated that the pseudo-Euclidean geometry is inherent in the fields associated with these interactions.

However, as a result of this discovery, the principle of relativity lost its fundamental role and was transformed into a partial implication claiming that all physical processes occur in a space-time having a pseudo-Euclidean geometry. Thus, the geometry of space-time began to play the fundamental role. The principle of relativity is a statement about the existence of a class of inertial reference systems in which all physical processes occur in the same way. In mathematical language this means that the equations describing the physical processes are forminvariant with respect to Lorentz transformations. At the limit $v/c \to 0$, the Lorentz transformations become the Galilean ones which provide the forminvariance of the equations of Newtonian mechanics.

But, as has been shown in [38-39], the statement that all physical processes occur in a space-time which has a pseudo-Euclidean geometry is much richer than the principle of relativity, in that it

allows a generalized principle of relativity to be formulated which will be valid not only in an inertial but also in a non-inertial coordinate system. In this connection, it must be noted that the literature not infrequently contains statements claiming that the special theory of relativity only deals with the description of phenomena in the inertial reference frames, while the description of phenomena in non-inertial reference frames is the prerogative of the GTR.

These statements are wrong. It follows, trivially, from Minkowski's fundamental discovery of the pseudo-Euclidean nature of space-time geometry that we may use any class of the admissible reference frames, both inertial and non-inertial, to describe the physical phenomena. The curvature tensor of this space-time, which defines its whole inertial geometry, vanishes both for the inertial and non-inertial reference frames. Because of this, it is quite conceivable to describe the physical phenomena either by the special theory of relativity or within non-inertial reference frames. This point was transparently clear to Fock [13].

Before formulating the generalized principle of relativity, let us briefly review the difference between general covariance and forminvariance.

Since the definitions of these two notions are quite close, covariance is frequently used to mean forminvariance when it is not implied by the former.

An equation is covariant under a certain transformation of coordinates, if the new unknown functions entering it, being expressed in terms of the new variables, satisfy the equations of the same form as the old functions in terms of the old variables. Thus, equation covariance does not reflect any physical principle but is a mathematical requirement.

As has been shown by Fock [13], for an equation to be covariant it is necessary that it is transformed according to a tensor law for any arbitrary, admissible coordinate transformation. We shall clarify this point with an example. The equations of relativistic mechanics,

$$\frac{Du^i(x)}{Ds} = F^i(x) \tag{1.39}$$

are covariant, since the new functions expressed in terms of the new variables $u'^i(x')$ satisfy an equation of the same form as the initial one (1.39) because of their tensor character:

$$\frac{D'u'^i(x')}{Ds'} = F'^i(x'),$$

given an arbitrary admissible coordinate transformation,

$$x'^i = x'^i(x), \tag{1.40}$$

i.e. having passed from the coordinates x to coordinates x', all the quantities in equation (1.39) are replaced with the appropriate primed quantities.

It should be specially emphasized here that the functional dependence of the metric space-time tensor g'_{ni} on the new coordinates under transformations (1.40) can in general be altered. This means that, while the metric tensor g_{ni} was a certain function of the coordinates x in the initial reference frame, then it might be an entirely different function of the coordinates x' in the primed coordinate frame. Since covariant equations always contain the space-time metric tensor or its derivatives, the functional dependence (a functional form) of the covariant equations on the new coordinates under transformation (1.40) will generally be altered.

This can be easily shown if one takes into account that, under the coordinate transformation (1.40), the metric tensor of space-time was transformed thus:

$$g'_{ni}(x') = \frac{\partial x^l}{\partial x'^n} \frac{\partial x^m}{\partial x'^i} g_{lm}(x(x')).$$

It is quite natural that the functional dependence of the covariant equations on the new coordinates under transformation (1.40) is not conserved: therefore, the description of phenomena in different coordinate systems will be different, i.e. the same phenomena in the different coordinate systems will generally occur in the different ways.

The requirement of forminvariance of the metric tensor under certain coordinate transformations (i.e. an invariability of the functional dependence of the metric tensor under this transformation) is more stringent than the requirement that equations are covariant. This requirement limits the class of coordinate systems to systems which transform into each other leaving the functional form of the metric space-time tensor unaltered, since the functional dependence of the tensor g_{ni} on the coordinates x in one reference frame is the same as the dependence of the tensor g'_{ni} on the coordinates x' in any other reference frame in this class.

Yet this requirement guarantees that, for the whole group of transformations leaving the metric forminvariant, the functional dependence of the field equations on the new coordinates is unaltered. As a result, all physical phenomena occur in the same way in all reference frames, which can be transformed with an unaltered metric, so that the system in which we are located cannot be identified.

Thus, covariance and forminvariance are quite different concepts. Transformations that ensure the covariance of field equations generally include transformations between different reference frames which are admissible but inadequate for the description of physical phenomena. By contrast, transformations which provide for the forminvariance of the metric space-time tensor (and hence the form-

invariance of the equations) involve only transformation between the reference systems which are equivalent from a physical standpoint, since all physical phenomena occur in the same way in these reference frames, given appropriate initial and boundary conditions.

Since space-time geometry does not alter and remains pseudo-Euclidean during the transition between the different reference frames, then for any reference frame, both inertial and non-inertial, a ten-parameter group exists for coordinate transformation that leaves the metric forminvariant. Thus, we are able to indicate, for any reference frame in the pseudo-Eucledian space-time, an infinite set of other reference frames mutual transformations between which leave the metric forminvariant.

This means that the generalized principle of relativity formulated in [38-39] is valid in the pseudo-Euclidean space-time: since whatever physical reference frame we choose (inertial or non-inertial), it is always possible to indicate an infinite set of other reference frames such that all physical phenomena occur in them in the same way they do in the initial reference frame. Thus we do not have nor can have any means to identify experimentally a specific reference frame from the infinite set.

Thus, Minkowskian geometry has a universal character as an inherent geometry for all known fields. Pseudo-Euclidean space-time is not a priori, i.e. given from the start, or having an independent existence. It is an integral part of the existence of matter, since it is the geometry by which matter is transformed.

The character of the space-time geometry greatly influences whether conservation laws for a closed system of interacting fields can be obtained.

Mathematically, the energy-momentum and angular momentum conservation laws reflect specific properties of space-time, i.e. its homogeneity and isotropy. Three sorts of space [40] exist that are sufficiently homogeneous and isotropic to allow the ten integrals of motion for a close system to be introduced, viz. a space with constant negative curvature (the Lobachevski space), a space with zero curvature (the pseudo-Euclidean space), and a space with constant positive curvature (the Riemannian space). The first two spaces are infinite (having an infinite volume) and the third space is closed (having a finite volume but not boundaries).

1.3 FUNDAMENTALS OF THE FIELD APPROACH TO THE DESCRIPTION OF GRAVITATIONAL INTERACTIONS

In order to treat a gravitational field as a physical field in the Faraday-Maxwell spirit having the common properties of being an energy-momentum carrier, it is sufficient to choose one of

the above geometries as the inherent one for gravitational fields. The experimental evidence from studies of strong, weak, and electromagnetic interactions indicates that the inherent space-time geometry for the fields associated with these interactions is pseudo-Euclidean. So at least for the time being we can assume this is the unified inherent geometry for all processes including gravitational.

This statement is one of the fundamental issues of the field approach we have developed as applied to the theory of gravitational interaction. It is quite evident that it results in all conservation laws of energy-momentum and angular momentum being satisfied, i.e. all ten integrals of motion exist for a system composed of gravitational and other fields. A gravitational field is, according to the field approach, like all other physical fields and is characterized by its own energy-momentum tensor which contributes to the complete energy-momentum tensor of the system. This is the main feature of our approach as compared with Einstein's GTR. It should be noted also that the integration of tensor quantities in a pseudo-Euclidean space-time has a definite meaning, as well as a general simplicity.

Another key question for the gravitational field theory is how exactly a gravitational field interacts with matter. When matter is acted on by gravity, the field can effectively alter its geometry, if the field is contained in the higher derivatives of the equation for motion of matter. The motion of material bodies and other physical fields under the influence of gravity will then be indistinguishable in pseudo-Euclidean space-time from their motions in an effective Riemannian space-time. The experimental evidence indicates the universal nature of the effect of gravity on matter, hence an effective Riemannian space-time will be unified for whole matter.

This leads us to what will be called the identity principle (geometrization principle). It is defined thus: The equations of motion for matter under the influence of gravitational field, in the pseudo-Euclidean space-time that has a metric tensor γ_{ni}, are identical to the equations of motion for matter in an effective Riemannian space-time with a metric tensor g_{ni} that is dependent on a gravitational field and on the metric tensor γ_{ni}.

This principle was introduced and formulated by us in [28], although essentially it had been suggested in [3]. It means that motion of matter in pseudo-Euclidean space-time under the influence of a gravitational field is described in the same way as the motions of matter in an appropriate, effective Riemannian space-time. With such an approach, gravitational field (treated as a physical one) is as if eliminated, when describing matter motion, and its energy, speaking figuratively, is spent to form the effective Riemannian space-time. Thus, the effective Riemannian space-time is a special energy-momentum carrier and by the identity principle, it consumes such an amount of energy which is contained in gravitational

field. Hence the propagation of the curvature waves in the Riemannian space-time takes into account the conventional transfer of energy by gravitational waves in the pseudo-Euclidean space-time. This means that using the field approach the curvature waves in a Riemannian space-time are the result of the existence of gravitational waves (in a spirit of Faraday-Maxwell) possessing the energy-momentum density.

It should be emphasized that the identity principle does not result from any physical principles. It is an adequate, independent principle which, on the one hand, ensures descriptions of motion of matter are equivalent and, on the other hand, governs the way a gravitational field interacts with matter, which corresponds to a definite choice of the Lagrangian density for the interaction. In particular, it also takes into account the physical fact that the inertial mass of the point body is equal to its gravitational mass.

Note also that, when introducing the geometrization principle, Einstein's great idea about Riemannian space-time geometry for matter is preserved. Yet this does not mean that we must return inevitably to the GTR. The GTR is just one implementation of this idea rather than vice versa. Therefore, when the idea that a gravitational field is a physical one and carries energy is combined with the identity principle we obtain gravitational field equations which differ from the Einstein equations, and which alter our concept of space-time and gravity. Since it is another implementation of Einstein's idea, this new theory of gravitation can describe all the gravitational experiments available, satisfies the correspondence principle and is accompanied by a number of fundamental corollaries.

It should be emphasized that the field approach to the theory of gravitational interaction does not, a priori, specify the nature of a gravitational field as we do not know what a real gravitational field is. It is possible, for example, that spin tensors or, perhaps, a scalar field may be needed to describe one adequately. For now, the lack of experimental material pertaining to gravitation opens a vast area to theoretical speculation, whence only time and new experimental facts will dictate a final choice of theory variant to be made.

1.4 A SYMMETRIC TENSOR GRAVITATIONAL FIELD

One possible implementation of the field approach is to use a symmetric tensor of rank two to describe a gravitational field. It should be noted that many authors have attempted to formulate a theory of gravitation in flat space-time using different fields, scalar, vector, and symmetric tensor, to do this. These attempts were yet fortuitous and did not have any clearly formulated field-theoretical requirements for their gravitation theories. As a result, the simplest

variants proposed in [41-71] either contradicted the experimental data available, or failed to ensure logical continuity and required the additional conditions to provide a positive definiteness for the energy of gravitational waves [72].

This point gave grounds for Thirring [73], and later other authors [74, 75, 127], to state that any way of constructing a theory of gravitation using flat space-time and assuming a Faraday-Maxwell gravitational field will inevitably lead to Einstein's GTR. Yet our analysis of Einstein's theory [1-8] as well as the search of other ways to construct a theory of gravitation [28-31] have shown that these statements are utterly without foundation. Einstein's theory, on the one hand, departed from the notion that gravitational fields possess energy-momentum and introduced a new type of field, i.e. one characterized by a curvature tensor, and on the other hand, Einstein's theory appears to be devoid of a fundamental principle, i.e. that the energy-momentum of matter and a gravitational field is conserved when the two are taken in conjunction. It is a too high price to pay to explain a small number of gravitational experiments.

For this reason, it is essential to compare the different classes of gravitation theories which conventionally use the symmetric tensor field of rank two, and to clear up which of them, introduces a gravitational field in the most suitable fashion from the physical standpoint.

When constructing a theory of gravitation, the key issue is the choice of the geometry inherent for a gravitational field. A flat space-time geometry is inherent in linear theories, and gravitation theories with linear equations for the free gravitational field are formulated in terms of flat space-time having a metric tensor γ_{ni}. We shall call these gravitation theories which are formulated in terms of flat space-time class (A) theories. Class (A) theories may also be nonlinear but it is essential that this nonlinearity does not appear in the higher-order derivatives of the field equations and thus does not alter the inherent geometry of space-time. Class (A) theories therefore make up a unified flat space-time, and so ensure that all ten conservation laws for a closed system are satisfied. Riemannian space-time itself, in which the motion of matter is described, is an effective one resulting from the action of the gravitational field φ_{ni} on matter.

Among the class (A) theories a subclass of bimetric theories should be noted for which the gravitational field tensor φ_{ni} is combined with the metric tensor γ_{ni} to form a new field variable in the density of the gravitational field Lagrangian L_g. For example, the metric tensor g_{ni} of the effective Riemannian space-time, in terms of which the equations of matter motion are formulated, and thereby the inherent geometry in this field variable is pseudo-Euclidean:

$$L = L_g\,(\gamma_{ni},\; g_{ni}\,(\gamma_{mk},\; \varphi_{mk})) + L_M\,(g_{ni},\; \varphi_A).$$

Rosen's theory may be taken as an example of the nonlinear theories of this subclass [65] and has the Lagrangian density

$$L_g = \frac{\sqrt{-\gamma}}{64\pi} \gamma^{ik} g^{nm} g^{pl} \left[D_i g_{nl} D_k g_{mp} - \frac{1}{2} D_i g_{nm} D_k g_{pl} \right],$$

where γ is the determinant of the metric tensor of the flat space-time, D_i the covariant derivative in the flat space-time; henceforth we assume $G = c = 1$.

In the bimetric theories the gravitational field φ_{ni} is in fact absent, since as a field variable we take its metric tensor g_{ni}. Therefore, a sufficiently solid physical substantiation of the coupling between the effective Riemannian space-time and the unified flat space-time exists here.

There are actually two physical space-times among the class (A) theories, viz. a flat space-time having the metric tensor γ_{ni}, in which the gravitational field equations are formulated, and a non-Euclidean space-time having the metric tensor g_{ni}, in which the motion of matter is stated. Both of these space-times are observable in reality. A gravitational wavefront moves along the geodesics of the flat space-time and hence gravitational waves can be used to identify the geometry of the pseudo-Euclidean space-time. An electromagnetic wavefront moves along the geodesics of the effective Riemannian space-time and hence electromagnetic waves and massive particles can be used to identify the geometry of this Riemannian space-time.

Since a nonlinear theory of the tensor field φ_{ni} will have nonlinear terms in the convolution of the derivatives in the expression for the Lagrangian density (in the higher-order derivatives of the field equations), a non-Euclidean space-time with an effective metric tensor $g_{ni} = g_{ni}(\gamma_{mk}, \varphi_{mk})$ is inherent in such a theory. We shall refer to theories of gravitation, which are formulated in terms of the effective Riemannian space-time, as class (B). The Lagrangian density of these theories has the form

$$L = L_g(g_{ni} + \varphi_{ni}) + L_M(g_{ni}, \varphi_A).$$

These theories deserve a special analysis.

Because flat space-time is nonobservable in the theories of this class, evidently, an adequate substantiation of the relation $g_{ni} = g_{ni}(\gamma_{mk}, \varphi_{mk})$ between the unified Riemannian space-time and the gravitational field φ_{ni} is missing in them. The unified Riemannian space-time in the theories of that class arises on the basis of the gravitational field φ_{ni} and of the nonobservable flat space-time. It should also be noted that the gravitational field equations in class (B) theories are, by necessity, nonlinear.

The subclass of the class (B) geometrized theories comprises a multitude of theories that have a complete geometrization, which

means that the Lagrangian density of a gravitational field depends only on the metric tensor g_{ni}:

$$L = L_g (g_{ni}) + L_M (n_{ni}, \varphi_A).$$

Einstein's theory belongs to this subclass of theories and corresponds to a specific choice of the Lagrangian density in the form $L_g = \sqrt{-g}R$. In theories with complete geometrization, flat space-time is entirely excluded from the description of motion both of matter and gravitational field. Neither the gravitational field φ_{ni}, nor the metric tensor γ_{ni} appear anywhere in the theory. The quantities g_{ni}, thereby, have two meanings as variables of the physical field and of the metric space-time tensor. As a result, gravitational field is not a field in the Faraday-Maxwell sense in the theories of this subclass.

It might be emphasized that class (A) and class (B) theories of gravitation differ substantially. No transformation of field variables or coordinates can convert one class of theory into the other.

Analysing the available possibilities thus, we arrive at the conlusion that only class (A) theories introduce gravitational field in a most suitable manner from the physical standpoint. This class of theories assumes that gravitation has a field in the spirit of Faraday-Maxwell and possesses all ten integrals of motion for a closed system of interacting fields. In the theories of this class effective Riemannian space-time used to describe the motion of matter takes into account the existence of a physical gravitational field and of a unified pseudo-Euclidean space-time.

As a result, we must, once again, make an immediate study of how to construct a gravitation theory which implements the field approach and which can describe gravitational interaction.

1.5 CONSERVATION LAWS FOR GRAVITATIONAL FIELD AND MATTER

Suppose we study the conservation laws for all the local theories of class (A) without limiting ourselves to a specific choice of Lagrangian density. Starting from the fundamentals of the field approach, the Lagrangian density of a system composed of matter and a gravitational field can be written as

$$L = L_g (\gamma_{ni}, \varphi_{ni}) + L_M (g_{ni}, \varphi_A) \tag{1.41}$$

for the theories of that class, where γ_{ni} is the metric tensor of the pseudo-Euclidean space-time having the signature $(+, -, -, -)$, φ_{ni} for the gravitational field, and φ_A for the remaining fields of matter.

Without the loss of generality, we can assume that the metric tensor g_{ni} of the Riemannian space-time is a local function dependent

on the metric tensor of the flat space-time, the gravitational field φ_{ni} and their partial derivatives up to the second order inclusively:

$$g_{ml} = g_{ml}(\gamma_{ni}, \partial_p\gamma_{ni}, \partial_{pk}\gamma_{ni}, \varphi_{ni}, \partial_p\varphi_{ni}, \partial_{pk}\varphi_{ni}), \quad (1.42)$$

where the following notation is introduced:

$$\partial_{nm}\varphi = \frac{\partial^2\varphi}{\partial x^n \partial x^m}.$$

The Lagrangian density L_M will be assumed to depend only on the fields φ_A, their first-order partial derivatives, and the metric tensor g_{ni}. It can be easily verified that here gravitational field's partial derivatives, up to the second order, will enter the Lagrangian density of matter.

The Lagrangian density of gravitational field will be assumed to depend on the metric tensor γ_{ni}, the gravitational field φ_{ni} and their partial derivatives, up to the third order inclusively.

To obtain the conservation laws, we shall use the covariant method of infinitesimal displacements.

Since the action J is a scalar, under an arbitrary, infinitesimal transformation of coordinates

$$x'^i = x^i + \xi^i(x) \quad (1.43)$$

the variation of the action of matter, δJ_M, and that of a gravitational field, δJ_g, will vanish.

Since both covariant and contravariant components of the metric tensor of the Riemannian space-time will enter the Lagrangian density of matter, we shall vary the Lagrangian density with respect to them as in the independent quantities and then will take into account the relations between their variations:

$$\delta g^{nm} = -g^{ni}g^{ml}\delta g_{il}.$$

In this way the density of the symmetric energy-momentum tensor of matter, T^{ni}, in the Riemannian space-time will have the form

$$T^{ni} = -2\frac{\Delta L_M}{\Delta g_{ni}} = -2\left(\frac{\delta L_M}{\delta g_{ni}} - g^{im}g^{nl}\frac{\delta L_M}{\delta g^{ml}}\right), \quad (1.44)$$

where $\frac{\delta L}{\delta\varphi}$ is the Euler-Lagrange variation:

$$\frac{\delta L}{\delta\varphi} = \frac{\partial L}{\partial\varphi} - \partial_n\left(\frac{\partial L}{\partial(\partial_n\varphi)}\right) + \partial_{nl}\left(\frac{\partial L}{\partial(\partial_{nl}\varphi)}\right) - \partial_{nli}\left(\frac{\partial L}{\partial(\partial_{nli}\varphi)}\right). \quad (1.45)$$

Completely analogously, we can vary the components γ_{ni} and γ^{in} of the metric tensor of flat space-time.

The variation of the action integral of matter under transformation (1.43) is written as

$$\delta J_{\mathrm{M}}=\int d^4x\left\{\frac{\Delta L_{\mathrm{M}}}{\Delta g_{ni}}\,\delta g_{ni}+\frac{\delta L_{\mathrm{M}}}{\delta\varphi_A}\,\delta\varphi_A+\mathrm{Div}\right\}=0, \tag{1.46}$$

where Div denotes the divergence terms which, when taken into account, result in relationships irrelevant to our analysis.

Introducing the notations

$$t^{nm}_{\mathrm{M}}=-2\,\frac{\Delta L_{\mathrm{M}}}{\Delta\gamma_{nm}}=-2\left(\frac{\delta L_{\mathrm{M}}}{\delta\gamma_{nm}}-\gamma^{ns}\gamma^{mk}\,\frac{\delta L_{\mathrm{M}}}{\delta\gamma^{sk}}\right), \tag{1.47}$$
$$t^{n}_{\mathrm{M}i}=\gamma_{im}t^{nm}_{\mathrm{M}}$$

for the energy-momentum tensor density of matter in flat space-time, the variation δJ_{M} of the action integral under the coordinate transformation (1.43) may also be written in another way equivalent to expression (1.46):

$$\delta J_{\mathrm{M}}=\int d^4x\left\{\frac{\delta L_{\mathrm{M}}}{\delta\varphi_{nm}}\,\delta\varphi_{nm}+\frac{\Delta L_{\mathrm{M}}}{\Delta\gamma_{nm}}\,\delta\gamma_{nm}+\frac{\delta L_{\mathrm{M}}}{\delta\varphi_A}\,\delta\varphi_A+\mathrm{Div}\right\}=0. \tag{1.48}$$

The variations $\delta\gamma_{nm}$, $\delta\varphi_{nm}$, $\delta\varphi_A$ and δg_{nm} under coordinate transformations (1.43) will have the form

$$\left.\begin{aligned}
\delta\gamma_{nm}&=-\gamma_{nl}D_m\xi^l-\gamma_{ml}D_n\xi^l,\\
\delta\varphi_{nm}&=-\varphi_{nl}D_m\xi^l-\varphi_{ml}D_n\xi^l-\xi^lD_l\varphi_{nm},\\
\delta\varphi_A&=-\xi^lD_l\varphi_A+F^{B;\,n}_{A;\,l}\varphi_BD_n\xi^l,\\
\delta g_{nm}&=-g_{nl}D_m\xi^l-g_{ml}D_n\xi^l-\xi^lD_lg_{nm}.
\end{aligned}\right\} \tag{1.49}$$

Allowing for these equalities, the variation of the action integral (1.48) may be written as follows:

$$\delta J_{\mathrm{M}}=\int d^4x\left\{\xi^l\left[2D_n\left(\frac{\delta L_{\mathrm{M}}}{\delta\varphi_{nm}}\,\varphi_{ml}\right)-D_nt^{n}_{\mathrm{M}l}-\frac{\delta L_{\mathrm{M}}}{\delta\varphi_{nm}}\,D_l\varphi_{nm}\right.\right.$$
$$\left.\left.-D_n\left(\frac{\delta L_{\mathrm{M}}}{\delta\varphi_A}\,F^{B;\,n}_{A;\,l}\varphi_B\right)-\frac{\delta L_{\mathrm{M}}}{\delta\varphi_A}\,D_l\varphi_A\right]+\mathrm{Div}\right\}=0. \tag{1.50}$$

From the arbitrariness of the displacement vector ξ^l in expression (1.50) an identity follows

$$D_lt^{l}_{\mathrm{M}n}-2D_l\left(\frac{\delta L_{\mathrm{M}}}{\delta\varphi_{lm}}\,\varphi_{mn}\right)+\frac{\delta L_{\mathrm{M}}}{\delta\varphi_{lm}}\,D_n\varphi_{lm}$$
$$+D_l\left(\frac{\delta L_{\mathrm{M}}}{\delta\varphi_A}\,F^{B;\,l}_{A;\,n}\varphi_B\right)+\frac{\delta L_{\mathrm{M}}}{\delta\varphi_A}\,D_n\varphi_A\equiv 0. \tag{1.51}$$

Another important identity can be obtained after substituting relation (1.49) into expression (1.46):

$$D_l\left(g_{nm}T^{lm}\right)-\frac{1}{2}\,T^{lm}D_ng_{ml}$$
$$=-D_l\left(\frac{\delta L_{\mathrm{M}}}{\delta\varphi_A}\,F^{B;\,l}_{A;\,n}\varphi_B\right)-\frac{\delta L_{\mathrm{M}}}{\delta\varphi_A}\,D_n\varphi_A. \tag{1.52}$$

Let us now express the covariant derivatives on the left-hand side of identity (1.52) in terms of partial derivatives and the connection γ^i_{nl} of the flat space-time. Taking into account that T^{ni} is the density of the tensor of weight 1, we shall obtain

$$\partial_i (g_{nm}T^{mi}) - \frac{1}{2} T^{im}\partial_n g_{im} = -D_i\left(\frac{\delta L_M}{\delta\varphi_A} F^{B;\,i}_{A;\,n}\varphi_B\right) - \frac{\delta L_M}{\delta\varphi_A} D_n\varphi_A.$$

However, the left-hand side of this expression is a covariant divergence in the Riemannian space-time with respect to the energy-momentum tensor density T^i_n of matter:

$$\partial_i (g_{nm}T^{mi}) - \frac{1}{2} T^{im}\partial_n g_{im} = \partial_i T^i_n - \Gamma^m_{ni}T^i_m = \nabla_i T^i_n = g_{nm}\nabla_i T^{im},$$

where, as usual, Γ^m_{ni} denotes the connection of the Riemannian space-time.

Hence, relation (1.52) will take the form

$$g_{ni}\nabla_m T^{im} = -D_i\left(\frac{\delta L_M}{\delta\varphi_A} F^{B;\,i}_{A;\,n}\varphi_B\right) - \frac{\delta L_M}{\delta\varphi_A} D_n\varphi_A.$$

Subtracting this equality from expression (1.51), we obtain

$$D_i t^i_{Mn} - 2D_i\left(\frac{\delta L_M}{\delta\varphi_{im}}\varphi_{mn}\right) + \frac{\delta L_M}{\delta\varphi_{im}} D_n\varphi_{im} = g_{ni}\nabla_m T^{im}. \tag{1.53}$$

It should be emphasized that this identity is valid regardless of whether the equations of motion of matter and of gravitational field are satisfied.

Likewise, it follows from the fact that the action of the gravitational field is invariant under transformation (1.43) that

$$D_i t^i_{gn} - 2D_i\left(\frac{\delta L_g}{\delta\varphi_{im}}\varphi_{mn}\right) + \frac{\delta L_g}{\delta\varphi_{im}} D_n\varphi_{im} = 0. \tag{1.54}$$

For the density t^i_{gn} of the symmetric energy-momentum tensor of the gravitational field we have, in the usual fashion

$$t^i_{gn} = -2\gamma_{nm}\frac{\Delta L_g}{\Delta\gamma_{im}}. \tag{1.55}$$

It follows from relations (1.53) and (1.54) that

$$D_i (t^i_{Mn} + t^i_{gn}) - 2D_i\left(\frac{\delta L}{\delta\varphi_{im}}\varphi_{mn}\right) + \frac{\delta L}{\delta\varphi_{im}} D_n\varphi_{im} = \nabla_i T^i_n. \tag{1.56}$$

Provided the gravitational field equations are satisfied,

$$\frac{\delta L}{\delta\varphi_{im}} = \frac{\delta L_g}{\delta\varphi_{im}} + \frac{\delta L_M}{\delta\varphi_{im}} = 0, \tag{1.57}$$

expression (1.56) can simplify to

$$D_i\,(t^i_{\mathrm{M}n}+t^i_{gn})=g_{ni}\nabla_m T^{im}. \tag{1.58}$$

This equality is a manifestation of the identity principle. It suggests that the covariant divergence of the sum of the densities of the energy-momentum tensors of matter and gravitational field in pseudo-Euclidean space-time was transformed into a covariant divergence of the density of the energy-momentum tensor of matter only in Riemannian space-time. These are just different ways of writing one and the same expression.

Provided the motion equations of matter are satisfied,

$$\frac{\delta L_{\mathrm{M}}}{\delta\varphi_A}=0, \tag{1.59}$$

expression (1.51) is simplified to

$$D_i t^i_{\mathrm{M}n}-2D_i\left(\frac{\delta L_{\mathrm{M}}}{\delta\varphi_{im}}\,\varphi_{mn}\right)+\frac{\delta L_{\mathrm{M}}}{\delta\varphi_{im}}\,D_n\varphi_{im}=0, \tag{1.60}$$

and from relation (1.52) the covariant conservation equation in Riemannian space-time follows automatically:

$$\nabla_n T^{ni}=\partial_n T^{ni}+\Gamma^i_{nm}T^{nm}=0. \tag{1.61}$$

This statement is common for theories with a geometrized density of the matter Lagrangian and is not associated with any particular variant of gravitation theory.

Further we see that from relations (1.60) and (1.54), provided the gravitational field equations (1.57) are satisfied, a covariant conservation law follows for the density of the complete symmetric energy-momentum tensor in pseudo-Euclidean space-time:

$$D_i\,(t^i_{\mathrm{M}n}+t^i_{gn})=0. \tag{1.62}$$

Thus, starting from the Lagrangian formalism, we have obtained an energy-momentum conservation law of matter and gravitational field in pseudo-Euclidean space-time. This fundamental law of nature reveals that in the field theory of gravitation no processes take place without energy-momentum being conserved. It also follows from expression (1.62) that a gravitational field within a pseudo-Euclidean space-time framework behaves like all other physical fields. It possesses energy-momentum and contributes to the density of the complete energy-momentum tensor of the system.

Starting from equality (1.62) and identity (1.58), we obtain

$$D_i\,(t^i_{\mathrm{M}n}+t^i_{gn})=g_{nm}\nabla_i T^{im}=0.$$

Consequently, given that the gravitational field equations (1.57) and the equations of motion of matter (1.59) are satisfied, the conservation law for the density of the complete energy-momentum ten-

sor (1.62) and the conservation law in the form (1.61) are simply the different ways of writing the same conservation law. Conservation law (1.62) manifests the fact that the density of the complete energy-momentum tensor of a system, composed of matter and a gravitational field, is conserved in pseudo-Euclidean space-time. This law has the common form of a conservation law. In Riemannian space-time the conservation law (1.61) is not a conservation law in the common sense, as the density T^{ni} of the energy-momentum tensor of matter need not be conserved,

$$\partial_n T^{ni} \neq 0.$$

As has been indicated by Einstein [12g]: "...Physically, the occurrence of the second term on the left-hand side shows that laws of conservation of momentum and energy do not apply in the strict sense for matter alone, or else that they apply only when the g_{ni} are constant, i.e. when the field intensities of gravitation vanish. This second term is an expression for momentum, and for energy, as transferred per unit of volume and time from gravitational field to matter."

In this case the second term in (1.61) expresses what effect a gravitational field has on the matter and shows that matter gains energy as if it were stored in a Riemannian geometry. Yet the gravitational field energy in this case was as if spent to build the Riemannian geometry. But which quantity is conserved is not seen from expression (1.61).

The absence of authentical conservation laws is peculiar to the whole subclass of gravitational theories that have a complete geometrization rather than just Einstein's theory. The Lagrangian density L_g of the gravitational field in the theories ~~of~~ this subclass depends on the field φ_{ni} and on the metric tensor γ_{ni} only via the metric tensor g_{ni} of Riemannian space-time. Hence in the theories of this subclass we have the following equation for the density of the symmetric energy-momentum tensor of matter and for a gravitational field in pseudo-Euclidean space-time:

$$\begin{aligned}
-\frac{1}{2} t^{ni} &= \frac{\Delta L}{\Delta \gamma_{ni}} = \frac{\Delta L_g}{\Delta \gamma_{ni}} + \frac{\Delta L_M}{\Delta \gamma_{ni}} = \frac{\Delta L}{\Delta g_{ml}} \frac{\partial g_{ml}}{\partial \gamma_{ni}} \\
&- \partial_p \left[\frac{\Delta L}{\Delta g_{lm}} \frac{\partial g_{lm}}{\partial (\partial_p \gamma_{ni})} \right] + \partial_{pq} \left[\frac{\Delta L}{\Delta g_{lm}} \frac{\partial g_{lm}}{\partial (\partial_{pq} \gamma_{ni})} \right] \\
&- \gamma^{is} \gamma^{np} \left\{ \frac{\Delta L}{\Delta g_{lm}} \frac{\partial g_{lm}}{\partial \gamma^{sp}} - \partial_q \left[\frac{\Delta L}{\Delta g_{lm}} \frac{\partial g_{lm}}{\partial (\partial_q \gamma^{sp})} \right. \right. \\
&\left. \left. - \partial_k \left(\frac{\Delta L}{\Delta g_{lm}} \frac{\partial g_{lm}}{\partial (\partial_{qk} \gamma^{sp})} \right) \right] \right\}.
\end{aligned}$$

Since in a geometrized theory the gravitational field equations have the form

$$\frac{\Delta L}{\Delta g_{lm}} = \frac{\delta L}{\delta g_{lm}} - g^{li} g^{mn} \frac{\delta L}{\delta g^{ni}} = 0,$$

the density of the symmetric energy-momentum tensor of matter and gravitational field vanishes in a pseudo-Euclidean space-time according to the gravitational field equations:

$$\frac{\Delta L}{\Delta \gamma_{ni}} = -\frac{1}{2} t^{ni} = 0. \tag{1.63}$$

An analogous conclusion about the ways the density of the symmetric energy-momentum tensor vanishes can also be obtained for the free gravitational field. Yet the equations of a free gravitational field have solutions with the nonzero curvature tensor R^i_{nlm}. Hence in theories with complete geometrization, the vanishing of the density of the energy-momentum tensor of the free gravitational field is not accompanied by the disappearance of the field φ_{ni} and, consequently, a dummy field exists which does not possess the energy-momentum density but brings about a curving of space-time, i.e. the formation of a Riemannian geometry. Whence it follows that a theory of a physical gravitational field possessing energy-momentum cannot in principle be reduced to the GTR on the basis of flat space-time. It follows therefore that statements a number of authors [74-75, 127] have made about the inevitability of such reduction are wrong. The subclass of gravitation theories with complete geometrization does not allow for the notion that a gravitational field can possess energy-momentum to be introduced.

To sum up, we have arrived at the following conclusions:

1. In local theories of class (A) a gravitational field described in a pseudo-Euclidean space-time is a physical field having energy-momentum. The motion of matter is described on the basis of an identity principle in effective Riemannian space-time which is produced by the energy-momentum of a gravitational field. The geometrical description arises in this approach by using field theory ideas for gravitational fields. This approach starts from the conservation laws.

2. In a subclass of theories with a complete geometrization, a gravitational field and matter have a unified geometry, but the gravitational field loses the properties of the physical field and does not possess any energy-momentum density. Within the framework of this approach, there are no field theory ideas about gravitational field in the Faraday-Maxwell sense.

The GTR implements one way of constructing the theory. It introduces a field of a new type, which is described by a curvature tensor and is not a Faraday-Maxwell field. Therefore, conservation laws are absent for matter and gravitational field taken in conjunction and, consequently, this theory does not satisfy the principle of correspondence to the Newtonian theory of gravitation.

1.6 A GAUGE-INVARIANT TENSOR FIELD

In this section every relation and equation will be formulated in terms of Cartesians, although they will certainly be covariant and written in an arbitrary coordinate system.

Consider the class (A) theories which have a Lagrangian density of the form (1.41). The equations of a gravitational field and those of the motion of matter have the form

$$\frac{\delta L_g}{\delta \varphi_{nm}} + \frac{\delta L_M}{\delta \varphi_{nm}} = 0. \tag{1.64}$$

$$\frac{\delta L}{\delta \varphi_A} = 0. \tag{1.65}$$

Among many theories that have the Lagrangian density (1.41), there are some in which the action integral is invariant with respect to the gauge transformation:

$$\varphi_{ni} \to \varphi_{ni} + \partial_i a_n + \partial_n a_i, \tag{1.66}$$

where a_i is an arbitrary gauge four-vector. Because the action integral of a free gravitational field is invariant under the gauge transformation (1.66), we have

$$\delta J_g = \int d^4x \left[-2a_n \partial_i \left(\frac{\delta L_g}{\delta \varphi_{ni}} \right) + \mathrm{Div} \right] = 0.$$

By virtue of the arbitrariness of the gauge vector a_n, we obtain

$$\partial_i \left(\frac{\delta L_g}{\delta \varphi_{ni}} \right) = 0.$$

From this equation and the field equation (1.64) there follows a conservation equation for a gravitational field source:

$$\partial_i \left(\frac{\delta L_M}{\delta \varphi_{ni}} \right) = 0.$$

As is known [76], the analogous conservation equations follow from the invariance of the Lagrangian density $L = L_A + L_M$ with respect to the gauge transformation of the vector potential $A_i \to A_i + \partial_i f$ in electrodynamics

$$\partial_i \frac{\delta L_A}{\delta A_i} = 0, \quad \partial_i \frac{\delta L_M}{\delta A_i} = 0.$$

Since the source is conserved in the field equations of gauge theory, it is usual to take the source in the equations of a gauge gravitational field theory to be a complete energy-momentum tensor of a system composed of "matter plus gravitational field". This results in the field equations becoming nonlinear and an assumption is usually advanced that the consecutive incorporation of these nonlinearities may result in the Einstein nonlinear theory of gravitation [74-75, 77, 127].

Yet in reality a similar hypothesis results, above all, in the gravitational field losing the property of being an energy-momentum carrier. Assuming the source

$$\frac{\delta L_{\mathrm{M}}}{\delta \varphi_{ni}}=\frac{1}{2}J^{ni}.$$

can be identified with the complete energy-momentum tensor

$$t^{ni}=t_{\mathrm{g}}^{ni}+t_{\mathrm{M}}^{ni}$$

immediately suggests that the energy-momentum tensor of the free gravitational field (at $L_{\mathrm{M}}=0$) vanishes. Such a theory does not incorporate the properties that are characteristic of other physical systems and for this reason we consider it to be inacceptable.

According to the Noether theorem, the invariance of the action integral relative to some transformation group leads to the existence of definitely conserved quantities. Invariance with respect to coordinate transformations is known to result in a conservation of energy-momentum tensor density t^{ni}. The invariance of the action integral with respect to the gauge transformations (1.66) results in the current J^{ni} being conserved. Since coordinate and gauge transformations are completely different, the quantities t^{ni} and J^{ni} are, naturally, completely different physical quantities.

The problem of constructing a guage-invariant tensor field theory lies mostly in constructing the conserved tensor current J^{ni}, or in other words, in constructing the Lagrangian density L_{M} of matter which will lead to the variation $\delta L_{\mathrm{M}}/\delta\varphi_{ni}$ being conserved. In order to solve this problem, it is necessary to analyse the spin states of the field described by a symmetric rank two tensor.

As has been shown in [78-79], the symmetric rank two tensor φ_{ni} may be given as a sum of the irreducible representations: one with spin 2, one with spin 1, and two with spin 0:

$$\varphi_{ni}=(P_2+P_1+P_0+P_{0'})_{ni}^{lm}\,\varphi_{lm}.$$

It is more convenient to use the momentum representation for the quantities P_S. Let us introduce the auxiliary operators

$$X_{ni}=\frac{1}{\sqrt{3}}\left(\gamma_{ni}-\frac{q_n q_i}{q^2}\right);\quad Y_{ni}=\frac{q_n q_i}{q^2},$$

by which the operators P_S can be given as

$$\left.\begin{aligned}
&P_0=X_{ni}X^{lm};\quad P_{0'}=Y_{ni}Y^{lm};\\
&P_1=(\sqrt{3}/2)\,(X_i^l Y_n^m+X_n^l Y_i^m+X_i^m Y_n^l+X_n^m Y_i^l);\\
&P_2=(3/2)\,(X_i^l X_n^m+X_n^l X_i^m)-X_{ni}X^{lm}.
\end{aligned}\right\}\qquad(1.67)$$

In the X-representation the projection operators, P_S, are nonlocal integro-differential ones:

$$P^{lm}_{ni}\varphi_{lm} = \int d^4y P^{lm}_{ni}(x, y)\,\varphi_{lm}(y)..$$

Using expressions (1.67), it is easy to verify that only the operators P_2 and P_0 can be conserved:

$$q_l P_{2\,ni}^{\;lm} = q_m P_{2\,ni}^{\;lm} = q_l P_{0\,ni}^{\;lm} = q_m P_{0\,ni}^{\;lm} = 0.$$

Hence, provided the field φ_{ni} enters the Lagrangian density L_g of the gravitational field simply as a combination, e.g.

$$f_{ni} = [(P_2 + \alpha P_0)\,\varphi]_{ni}, \tag{1.68}$$

the Lagrangian density, and consequently the equations of a free gravitational field as well, will be invariant with respect to the gauge transformation (1.66). Yet an application of expression (1.68) is not really convenient, since it is integro-differential and thus results in nonlocal field equations.

For the field equations to be local, we need a differential coupling between the fields f_{ni} and φ_{ni}. This can be gained by taking, for example, the following combination

$$f_{ni} = \square^2 [(P_2 + \alpha P_0)\,\varphi]_{ni}.$$

In this case, the tensor f_{ni} will be expressed in terms of the fourth derivatives of the field function φ_{ni}. Yet among all the values of α, the case where $\alpha = -2$ is distinct in that it allows f_{ni} to be put down without any fourth derivatives of the field φ_{ni} and using only a combination of second derivatives:

$$f_{ni} = \square\, [(P_2 - 2P_0)\,\varphi]_{ni}. \tag{1.69}$$

It can be seen that the operator $\square\,(P_2 - 2P_0)$ is gradient invariant and a local operator of the lowest possible order: in a theory using a symmetric rank two tensor field there is no other local operator that incorporates lower-order derivatives and results in gauge invariance. Thus we have

$$\begin{aligned} &f_{ni} = \square\theta_{ni} - \partial_i\partial^m\theta_{mn} - \partial_n\partial^m\theta_{mi} + \gamma_{ni}\partial^l\partial^m\theta_{lm};\\ &\partial^i f_{ni} = 0, \end{aligned} \tag{1.70}$$

where we introduced the notation

$$\theta_{lm} = \varphi_{lm} - \frac{1}{2}\gamma_{lm}\varphi^n_n. \tag{1.71}$$

Here the vector field and the spin 0′ field vary under gradient transformation (1.66) and are eliminated from the theory.

Since the whole theory has to be gauge invariant, we assume that the fields φ_{lm} only enter the coupling equations $g_{ni} = g_{ni}(\varphi_{lm})$

via the field f_{lm}. Moreover, we assume that the metric tensor g_{ni} of Riemannian space-time is a local function of the metric tensor of flat space-time and only the fields f_{lm}. We will not now make any suggestion to the form of this function except to require the quadratic form with the coefficients $g_{\alpha\beta}$ to be negative definite, the component g_{00} being a positive quantity. Thus the parameter x^0 will have the dimensions of time and the parameters x^α, those of the space coordinate in Riemannian space-time.

1.7 THE GRAVITATIONAL FIELD EQUATIONS IN THE FIELD THEORY OF GRAVITATION

A gauge invariant theory that leads to linear equations for a free gravitational field is the simplest variant of all the class (A) theories. From now on we will call this theory of gravitation the field theory of gravitation. The Lagrangian density of a gravitational field in the field theory does not incorporate derivatives of the fields f_{nm} higher than the first order and may be written in its most general form as

$$L_g = \frac{1}{64\pi}\{\partial_i f_{nm}\partial^i f^{mn} - b\partial_i f\partial^i f - m_g^2[\alpha f_{nm}f^{nm} + \beta f^2]\},$$

where $f = f^n_n$.

If $\alpha \neq 0$ or $\beta \neq 0$, then the equations obtained will describe a gravitational field, a quantum of which (the graviton) has a nonzero rest mass. However, we expect a gravitational wavefront to propagate at the fundamental velocity $v = c$, and so a graviton rest mass should vanish thus making it necessary to put $\alpha = \beta = 0$.

By trying various values of b, we can realize different physical situations. The energy of a free gravitational field may be shown to be a positive quantity once $b \leqslant 1/2$. Moreover, when $b < 1/2$ the scalar component of gravitational waves is emitted, the value of this component and its energy being substantially dependent on the value of b inasmuch as at $b < 1/2$ the scalar component transfers positive energy. Later on we will not need this degree of generality since we can assume gravitational waves (gravitons) to be characterized by the spin value $s = 2$ and by a positive definite energy. For this reason, we set $b = 1/2$ from now on in order to rule out scalar component emission.

Thus, we arrive at the Lagrangian density of a free gravitational field:

$$L_g = \frac{1}{64\pi}\left\{\partial_i f_{nm}\partial^i f^{nm} - \frac{1}{2}\partial_i f\partial^i f\right\}. \tag{1.72}$$

This Lagrangian density is the simplest possible that is invariant under the gauge transformations (1.66) of the fields φ_{ni}. The fields f_{nm} are also liable to the gauge transformation

$$f_{nm} \to f_{nm} + \partial_n a_m + \partial_m a_n - \gamma_{nm}\partial_i a^i, \tag{1.73}$$

and the conditions $\partial^n f_{nm} = 0$ are not violated if the gauge vectors a^n satisfy the homogeneous equations $\Box a^n = 0$.

It must be stressed that the symmetric fields f_{nm} are not independent by virtue of the four conditions, $\partial^n f_{nm} = 0$, they satisfy, and so when deriving the field equations, the Euler-Lagrange variation has to be taken over the field φ_{ni}, as this is the only field for which all ten components are independent. Yet once the field equations have been defined the Euler-Lagrange variation is taken over the f_{nm} field, the account should be taken of the four auxiliary conditions $\partial^n f_{nm} = 0$ as well. These conditions are satisfied by the field, i.e. we must state a problem about the conditional extremum. In both cases we shall arrive at equivalent field equations. Note also that the variation of the Lagrangian density of matter may be obtained in two ways, either explicitly by putting down the Euler-Lagrange variation (1.45) over the field φ_{ni}, or by using the fact that the gravitational field only enters the Lagrangian density of matter through the field f_{nm} (1.70), which, in turn, only comes into the Lagrangian density through the Riemannian space-time metric tensor. Both ways yield the same result.

Introducing the notation

$$h^{lm} = \frac{1}{2} T^{np} \frac{\partial g_{np}}{\partial f_{ik}} (\delta^l_i \delta^m_k + \delta^l_k \delta^m_i - \gamma_{ik}\gamma^{lm}) \tag{1.74}$$

and incorporating relation (1.70), the gravitational field equations (1.64) can be obtained as:

$$\Box^3\theta^{lm} - \partial^l\partial_n\Box^2\theta^{nm} - \partial^m\partial_n\Box^2\theta^{nl} + \gamma^{lm}\partial_n\partial_p\Box^2\theta^{np} = -16\pi J^{lm}, \tag{1.75}$$

where

$$J^{lm} = \Box h^{lm} - \partial^l\partial_n h^{nm} - \partial^m\partial_n h^{nl} + \gamma^{lm}\partial_n\partial_i h^{ni}.$$

Using definition (1.70), the field equations (1.75) take the form

$$\Box^2 f^{lm} = -16\pi J^{lm}. \tag{1.76}$$

It is easy to check that the gravitational field equations, both as (1.75) and as (1.76), are invariant under gauge transformations (1.66) with an arbitrary gauge vector a^n. If we take the full divergence with respect to one of the indices in the field equations (1.75) and (1.76), we get the identity $0 \equiv 0$. For this reason, even though the field φ_{ni} has ten independent components, the equation structure is such that the four components corresponding to spins 1 and 0 are automatically eliminated from the equations. As a result the

field equations retain only the six independent components corresponding to spins 0 and 2 for which we have six independent field equations, since the four conditions (1.70) are fulfilled by gauge invariance.

The gravitational field equations (1.76) may be simplified by using gauge transformation (1.66) and imposing additional conditions on the field functions. Because the choice of gauge is arbitrary, it becomes necessary when solving a concrete problem to make the gauge conditions explicit by some method, for example, by imposing additional conditions. The fact that in gauge theory the Euler-Lagrange variation satisfies the four identities $\partial_i \delta L/\delta\varphi_{im} = 0$ also means that to solve the field equations for a concrete problem it is necessary to impose at least four additional conditions on the field. Under gauge transformations (1.66) and by virtue of (1.71), the fields θ_{nm} are subjected to the gauge transformation

$$\theta_{nm} \to \theta_{nm} + \partial_n a_m + \partial_m a_n - \gamma_{nm}\partial_k a^k. \tag{1.77}$$

The most general additional conditions which are linear over the $\square^2\theta^{nm}$ field, are

$$\partial_n \square^2\theta^{nm} = A\partial^m \square^2\theta^n_n. \tag{1.78}$$

On fulfilling (1.78), the gravitational field equations can be written as

$$\square^3\theta^{lm} - 2A\partial^l\partial^m\square^2\theta^n_n + A\gamma^{ml}\square^3\theta^n_n = -16\pi J^{lm}.$$

It can be verified that given the additional conditions (1.78), the left-hand side of these equations can be conserved as well. At $A = 0$, the gravitational field equations acquire their simplest possible form:

$$\square^3\theta^{lm} = -16\pi J^{lm} \tag{1.79}$$

with additional conditions

$$\partial_n \square^2\theta^{nm} = 0. \tag{1.80}$$

Thus, the gravitational field equations are, in our case, those which have higher derivatives, and equations (1.79) are also invariant under the gauge transformations (1.77), whilst not violating the additional conditions (1.80).

If we introduce the field H^{nm} thus,

$$\square H^{nm} = h^{nm}, \tag{1.81}$$

equation (1.79) will take the form

$$\square^3\theta^{lm} = -16\pi\square\{h^{lm} - \partial^l\partial_n H^{nm} - \partial^m\partial_n H^{nl} + \gamma^{lm}\partial_n\partial_p H^{np}\}.$$

Since later on we shall only be interested in the causally conditioned solutions according to [80], we may "cancel" these equations by the D'Alembert operator. Introducing the notation

$$\psi_{nm} = \Box\theta_{nm} \tag{1.82}$$

for the causally conditioned solutions, the gravitational field equations will be obtained in the following form:

$$\Box\psi^{lm} = -16\pi\{h^{lm} - \partial^l\partial_n H^{nm} - \partial^m\partial_n H^{nl} + \gamma^{lm}\partial_n\partial_p H^{np}\}.$$

The tensor current on the right-hand side of this equation satisfies the following condition outside the source:

$$\Box\{h^{lm} - \partial^l\partial_n H^{nm} - \partial^m\partial_n H^{nl} + \gamma^{lm}\partial_n\partial_p H^{np}\} = 0.$$

Therefore, outside the matter, this tensor current may be eliminated during the gauge transformation. Indeed, since the additional conditions (1.80) allow transformations (1.77) with the gauge four-vector to satisfy

$$\Box^3 a^n = 0, \tag{1.83}$$

we are able to perform the following gauge transformation:

$$\psi^{nm} \to \psi^{nm} + \partial^n \Box a^m + \partial^m \Box a^n - \gamma^{nm}\partial_l \Box a^l. \tag{1.84}$$

Outside the source, as a gauge four-vector, we choose one that satisfies the condition

$$\Box^2 a^n = 16\pi\partial_m H^{mn}.$$

Since the equations $\Box H^{nm} = 0$ are fulfilled outside the source, the gauge four-vector satisfies equation (1.83), in this region too, and as a result the additional conditions (1.80) are automatically satisfied. The form of the gauge vector inside the source is not essential at present. After the gauge transformation (1.83) outside the source, we obtain the gravitational field equations as

$$\Box\psi^{nm} = 0.$$

This means that the tensor current

$$I^{nm} = h^{nm} - \partial^n\partial_l H^{lm} - \partial^l\partial^m H^n_l + \gamma^{nm}\partial_l\partial_p H^{pl} \tag{1.85}$$

differs from zero only outside the matter and so under the given gauge the gravitational field equations becomes

$$\Box\psi^{nm} = -16\pi I^{nm}. \tag{1.86}$$

These equations allow the gauge transformations (1.77) for the class of the vectors that satisfy

$$\Box^2 a^n = 0.$$

Therefore, we will solve equations (1.86) using the additional conditions that $\partial_l \psi^{ln} = 0$, which only leave the possibility of performing the gauge transformations for this class. This choice of additional condition is in line with the Fock theorem [13] in which the solution of the homogeneous wave equation $\Box\, \partial_l \psi^{ln} = 0$, bounded over the whole space and satisfying the Sommerfeld radiation condition, is identically equal to zero:

$$\partial_l \psi^{ln} = 0. \tag{1.87}$$

Thus, the gravitational field equation

$$\Box\, \psi^{nm} = -16\pi I^{nm} \tag{1.88}$$

is obtained using the additional conditions

$$\partial_n \psi^{nm} = 0. \tag{1.89}$$

Note further that, using the accepted notation for (1.82), the expression $\Box f_{nm}$ may be written in the form

$$\Box\, f_{nm} = \Box\, \psi_{nm} - \partial_n \partial^l \psi_{lm} - \partial_m \partial^l \psi_{ln} + \gamma_{nm} \partial^l \partial^p \psi_{lp}.$$

This expression is also invariant under transformations (1.84) with any gauge vector a^n, but in this case the operator $\Box f_{nm}$ will have the initial form. We may simplify this operator, the additional conditions (1.89) are satisfied within the accepted gauge, and get

$$\Box\, f_{nm} = \Box\, \psi_{nm}. \tag{1.90}$$

Note that the resultant operator $\Box f_{nm}$ is also invariant under the gauge transformations (1.84) and do not violate the additional conditions (1.89). Relations (1.90) allow the gravitational field equations to be rewritten as

$$\Box\, f^{nm} = -16\pi I^{nm} \tag{1.91}$$

using the additional conditions

$$\partial_n f^{nm} = 0. \tag{1.92}$$

It is stressed that the tensor current I^{nm} on the right-hand side of (1.91) is concentrated within the matter.

Note also that equations (1.91) of the field theory of gravitation may not only be formulated for an inertial, but also for a non-inertial coordinate system. In addition the field equations are forminvariant for the transition from one non-inertial coordinate system to another and for every infinite manifold of non-inertial coordinate system. As regards inertial coordinate systems, the field equations are Lorentz-invariant, when passing from one to another. This requires the relativity principle to be extended [30] and can be formulated thus: no physical phenomena, including gravitational, enable us

to find out whether we are at rest or in a state of a uniform and translational motion.

Let us emphasize that the relativity principle does not require the propagation velocity of an electromagnetic wavefront, i.e. the velocity of light to be constant. It is natural that during an interaction with external gravitational fields the velocity of light, as well as that of other bodies in motion, is not constant.

1.8 MINIMUM COUPLING EQUATIONS

To close the theoretical scheme, we will now indicate a coupling equation between the metric tensor g_{ni} of the effective Riemannian space-time and the gravitational field f_{ni}.

Since the choice of the coupling equation in the field theory of gravitation is equivalent to the choice of the density of the Lagrangian of an interaction between a gravitational field and the other fields of matter, the construction of the coupling equation will be analogous to the construction of the interaction Lagrangian density in the theories of other physical fields. Thus, for example, in electrodynamics a "minimum Lagrangian" is chosen as the interaction Lagrangian density.

It is reasonable in the field theory of gravitation, too, to choose as the coupling equation the minimum coupling equation that can describe all the available experiments in the case of weak gravitational fields. Using the linear approximation usually considered, the tensor current I^{nm} (1.85) has to be taken in the absence of gravitational field. Since the only physical and symmetric tensor of rank two which satisfies the conservation law in this approximation is the energy-momentum tensor of matter, we must require that the tensor current I^{nm} should become identical to the energy-momentum tensor of matter in the zero approximation for gravitational field:

$$I^{nm}(f_{lp}=0)=T^{nm}. \tag{1.93}$$

This identity requirement allows us to reconstruct the structure of the coupling equation, $g_{ni}=g_{ni}(f_{lm})$, unequivocally with the linear approximation. In fact, by taking advantage of expressions (1.74), (1.81), and (1.85), we see that the identity requirement (1.93) will result in the following coupling equation for the linear approximation:

$$g_{nm}=\gamma_{nm}+f_{nm}-\frac{1}{2}\gamma_{nm}f. \tag{1.94}$$

It might be assumed that relation (1.94) is the minimum coupling equation and is always satisfied and not only for the linear approximation over a weak field f_{nm}. However, a theory having this sort of coupling equation will then fall into class of so-called "quasilinear" theories of gravitation (according to Will's terminology)

But, as has been shown in [81], any "quasilinear", asymptotically Lorentz-invariant theory of gravitation is inconsistent with experimental results. Hence, relation (1.94) must only be an expansion of the minimum-coupling equation up to the linear terms in the weak field f_{nm}. Thus, the minimum-coupling equation should be quadratic in the field f_{nm}:

$$g_{nm}=\gamma_{nm}+f_{nm}-\gamma_{nm}f/2+[b_1f_{nl}f^l_m+b_2f_{nm}f+b_3\gamma_{nm}f_{il}f^{il} + b_4\gamma_{nm}f^2]/4, \tag{1.95}$$

with the as yet undefined minimum coupling parameters b_1, b_2, b_3, b_4.

As will be seen below, the condition that the post-Newtonian expressions for the inertial and gravitational mass of a static, spherically symmetric body should be identical implies the following relationship between the minimum coupling parameters: $2(b_1 + b_2 + b_3 + b_4) = 1$.

It might be possible to consider other, more complicated, coupling equations which would transform into the minimum-coupling equation (1.95) only with the weak field approximation, but there is no reason for this complication. The minimum coupling equation (1.95) describes every gravitational experiment.

Hence, all the subsequent analysis, too, will start from minimum-coupling equation (1.95). We will assume as a basic physical requirement singularities in the effective Riemannian space-time metric at finite values of the matter density inside a gravitational field source are absent. This is done in order to impose some restriction upon the values of minimum coupling parameters and to exclude in the field theory of gravitation the appearance of objects resembling black holes.

What is more, we shall require that the theory does not have Olbers type paradoxes in the description of a model of the Universe.

It should be noted that because of the minimum-coupling equation (1.95), the non-diagonal components of the metric tensor g_{nm} of the Riemannian space-time need not vanishing even if the non-diagonal components of the gravitational field f_{nm} vanish.

For the non-diagonal components of the tensor g_{nm} to vanish when the appropriate non-diagonal components of gravitational field vanish, it is necessary and sufficient to set $b_1 = 0$. In doing so we obtain the simplest minimum-coupling equation (S-M coupling):

$$g_{nm} = \gamma_{nm} + f_{nm} - \gamma_{nm}f/2 + [b_2f_{nm}f + b_3\gamma_{nm}f_{il}f^{il} + b_4\gamma_{nm}f^2]/4. \tag{1.96}$$

The condition that the post-Newtonian expressions for the gravitational and inertial mass of a static, spherically symmetric body coincide requires the S-M coupling parameters to satisfy the relationship

$$2(b_2 + b_3 + b_4) = 1.$$

1.9 CONSERVATION LAWS IN THE FIELD THEORY OF GRAVITATION

In Sec. 1.5 conservation laws were obtained that are valid for all class (A) gravitation theories. Given the differential conservation law (1.62) for the density of the complete symmetric energy-momentum tensor of a system in flat space-time for the theories in this class, the appropriate integral conservation law can be obtained.

In Cartesian coordinates we get

$$\partial_n [t_g^{ni} + t_M^{ni}] = 0. \tag{1.97}$$

Integrating this expression over a volume V at $i = 0$ and assuming the absence of mass flux through the surface enclosing this volume, we obtain

$$-\frac{\partial}{\partial t} \int dV [t_g^{00} + t_M^{00}] = \int dS_\alpha t_g^{0\alpha}. \tag{1.98}$$

Thus when gravitational waves are emitted, the energy of the source has to change and, if the gravitational waves transfer positive energy, the source energy has to decrease.

All those conclusions and relationships are valid for the field theory too as this is a concrete specimen of a class (A) theory. Since the symmetric and canonical energy-momentum tensors only differ by the divergence of the antisymmetric rank three tensor, conservation laws (1.62) and (1.97) hold for the canonical energy-momentum tensor as well.

The canonical energy-momentum tensor of a free gravitational field may be obtained in the following way. Let us start with the identity

$$\frac{\partial L_g}{\partial x^p} = \partial_n \left[\frac{\partial L_g}{\partial (\partial_n f_{lm})} \partial_p f_{lm} \right] - \partial_p f_{lm} \partial_n \left[\frac{\partial L_g}{\partial (\partial_n f_{lm})} \right]. \tag{1.99}$$

According to (1.91), a free gravitational field satisfies the equation

$$\partial_n \left[\frac{\partial L_g}{\partial (\partial_n f_{lm})} \right] = \Box f_{lm} = 0,$$

hence expression (1.99) represents the way the divergence of the canonical energy-momentum tensor of a free gravitational field vanishes. Whence

$$\tilde{t}^n_{gp} = - L_g \delta^n_p + \frac{\partial L_g}{\partial (\partial_n f_{lm})} \partial_p f_{lm}. \tag{1.100}$$

Using expression (1.72) for the Lagrangian density of a free gravitational field, we obtain

$$\tilde{t}^{n}_{gp}=\frac{1}{64\pi}\left\{-\delta^n_p\left[\partial_i f_{lm}\partial^i f^{lm}-\frac{1}{2}\partial_i f\partial^i f\right]\right.$$
$$\left.+2\partial_p f_{lm}\partial^n f^{lm}-\partial_p f\partial^n f\right\}. \quad (1.101)$$

In order to discover the symmetric energy-momentum tensor of the gravitational field t^{ni}_g, the Lagrangian density L_g of the gravitational field and an expression for f_{ni} have to be written in an explicit covariant form. Transforming expression (1.72) from Cartesian coordinate to an arbitrary curvilinear reference frame, we obtain

$$L_g=\frac{\sqrt{-\gamma}}{64\pi}\gamma^{ik}\left[\gamma^{ln}\gamma^{mp}-\frac{1}{2}\gamma^{lm}\gamma^{np}\right]D_i f_{lm}D_k f_{np}. \quad (1.102)$$

Likewise, from expression (1.69) we have

$$f_{ik}=\gamma^{lm}[D_l D_m\varphi_{ik}-D_i D_l\varphi_{mk}-D_k D_l\varphi_{mi}$$
$$+D_i D_k\varphi_{lm}+\gamma_{ik}\gamma^{np}(D_l D_n\varphi_{mp}-D_n D_p\varphi_{lm})]. \quad (1.103)$$

To reduce the length of subsequent expressions, let us introduce the following notation:

$$\Lambda^{ik}=-A^{lm}[\partial_l\partial_m\varphi^{ik}-\partial^i\partial_l\varphi^k_m-\partial^k\partial_l\varphi^i_m+\partial^i\partial^k\varphi_{lm}]$$
$$+fA^{ik}/2+A^n_n[f^{ik}-\gamma^{ik}f/2]+\partial_s\{\varphi^i_n[-\partial^s A^{kn}+2\partial^n A^{sk}$$
$$+2\gamma^{sk}\partial_l A^{ln}-\gamma^{kn}\partial_l A^{ls}-\partial^k A^{sn}+\gamma^{kn}\partial^s A^l_l$$
$$-2\gamma^{sk}\partial^n A^l_l]+\varphi^s_n[\partial^i A^{kn}-\partial^n A^{ik}-\gamma^{ik}\partial_l A^{ln}$$
$$+\gamma^{ik}\partial^n A^l_l]+2\gamma^{ks}A^{np}\partial^i\varphi_{np}-A^{sn}\partial^i\varphi^k_n-3A^{kn}\partial^i\varphi^s_n$$
$$+2A^{ks}\partial^i\varphi^n_n-\gamma^{ik}A^{nm}\partial^s\varphi_{nm}+3A^{kn}\partial^s\varphi^i_n-A^{ik}\partial^s\varphi^n_n$$
$$-2\gamma^{sk}A^{ln}\partial_l\varphi^i_n-2A^{sk}\partial_l\varphi^{li}+A^{ns}\partial_n\varphi^{ik}+\gamma^{ik}A^{ln}\partial_l\varphi^s_n$$
$$+A^{ik}\partial_n\varphi^{ns}+A^l_l[2\partial^i\varphi^{ks}-2\gamma^{ks}\partial^i\varphi^n_n-2\partial^s\varphi^{ik}$$
$$+\gamma^{ik}(\partial^s\varphi^n_n-\partial_n\varphi^{ns})+2\gamma^{ks}\partial_n\varphi^{ni}]\}/2. \quad (1.104)$$

The symmetric energy-momentum tensor of gravitational field may be obtained by substituting expressions (1.102) and (1.103) into relation (1.55). In Cartesian coordinates

$$t^{ik}_g=\frac{1}{64\pi}\left\{-\gamma^{ik}\left[\partial_l f_{nm}\partial^l f^{nm}-\frac{1}{2}\partial_n f\partial^n f\right]\right.$$
$$\left.+2\partial^i f_{nm}\partial^k f^{nm}-\partial^i f\partial^k f\right\}+\frac{1}{16\pi}\left\{\partial_l f^{ni}\partial^l f^k_n-\frac{1}{2}\partial_l f^{ik}\partial^l f\right\}$$

$$-\frac{1}{32\pi}\partial_l\{f_p^i[\partial^l f^{kp}+\partial^k f^{lp}]-f^{ik}\partial^l f+f_n^k[\partial^l f^{ni}+\partial^i f^{nl}]$$
$$-f^{nl}[\partial^i f_n^k+\partial^k f_n^i]\}-2\Lambda^{(ik)}, \quad \bullet \tag{1.105}$$

where, as usual, symmetrization is performed over the parenthesized indices:

$$\Lambda^{(ik)}=\frac{1}{2}(\Lambda^{ik}+\Lambda^{ki}).$$

Here the tensor A^{nm} from expression (1.104) has the following form:

$$A^{nm}=-\frac{1}{32\pi}\Box\left[f^{nm}-\frac{1}{2}\gamma^{nm}f\right].$$

Outside the matter $\Box f_{nm}=0$, hence the expression for t_g^{ik} is substantially simplified:

$$t_g^{ik}=\tilde{t}_g^{ik}+\frac{1}{32\pi}\partial_l\{f_n^l[\partial^i f^{kn}+\partial^k f^{ni}]-f_n^i\partial^k f^{nl}-f_n^k\partial^i f^{nl}\}, \tag{1.106}$$

where $\tilde{t}_g^{ik}$ is the canonical energy-momentum tensor of a free gravitational field (1.101).

We shall show that the symmetric energy-momentum tensor t_g^{ik} of a gravitational field differs, in the wave zone, from the canonical energy-momentum tensor $\tilde{t}_g^{ik}$ only by non-wave terms which diminish faster than $1/r^2$. Since the expression

$$f_{nm}=\frac{a_{nm}(t-r,\ \theta,\ \varphi)}{r}+O\left(\frac{1}{r^2}\right)$$

is valid in the wave zone, for an arbitrary function $F(f_{nm})$ we have

$$\partial_\alpha F=n_\alpha\frac{\partial}{\partial t}F+O\left(\frac{1}{r}F\right),$$

where

$$n_\alpha=\frac{x_\alpha}{r}.$$

Therefore expression (1.106) may be written down in the form

$$t_g^{ik}=\tilde{t}_g^{ik}+\frac{1}{32\pi}\frac{\partial}{\partial t}\{[f^{0l}+n_\alpha f^{\alpha l}][\partial^i f_l^k+\partial^k f_l^i]$$
$$-f_n^i\partial^k[f^{0n}+n_\alpha f^{\alpha n}]-f_l^k\partial^i[f^{0l}+n_\alpha f^{\alpha l}]\}+O\left(\frac{1}{r^3}\right).$$

Denoting differentiation over time by a dot, we have from the additional conditions (1.92):

$$\dot{f}^{0m}+n_\alpha\dot{f}^{\alpha m}=O\left(\frac{1}{r^2}\right). \tag{1.107}$$

Integrating this expression over time and assuming the integration constants vanish, since waves should not have a time-dependent component, we obtain

$$f^{0m}+n_\alpha f^{\alpha m}=O\left(\frac{1}{r^2}\right). \tag{1.108}$$

Whence it follows that in the wave zone the symmetric energy-momentum tensor of a gravitational field differs from the canonical energy-momentum tensor by a non-wave quantity which diminishes, with rising r, faster than $1/r^2$:

$$t_g^{ih}=\tilde{t}_g^{ih}+O\left(\frac{1}{r^3}\right). \tag{1.109}$$

Hence in the wave zone every calculation using either the symmetric or canonical energy-momentum tensor of a gravitational field will yield the same result. These tensors are also equivalent for computing the integral characteristics of gravitational radiation. In fact, we have from expression (1.106)

$$t_g^{00}=\tilde{t}_g^{00}+\frac{1}{16\pi}\,\partial_\alpha\{f^{\alpha l}\dot{f}_l^0-f_l^0\dot{f}^{\alpha l}\}.$$

Hence

$$\int t_g^{00}\,dV=\int \tilde{t}_g^{00}\,dV+\frac{1}{16\pi}\int dS_\alpha\,[f^{\alpha l}\dot{f}_l^0-f_l^0\dot{f}^{\alpha l}].$$

Once the boundary of the integration region is found in the wave zone, we have, by virtue of relations (1.107) and (1.108),

$$f^{\alpha l}\dot{f}_l^0-f_l^0\dot{f}^{\alpha l}=n_\beta\,[f^{\alpha l}\dot{f}_l^\beta-\dot{f}^{\alpha l}f_l^\beta]+O\left(\frac{1}{r^3}\right).$$

Choosing a sphere of radius r as the integration surface ($dS_\alpha=$ $=-r^2n_\alpha\,d\Omega$), we obtain

$$\int t_g^{00}dV=\int \tilde{t}_g^{00}dV+O\left(\frac{1}{r}\right). \tag{1.110}$$

Besides, it follows from relation (1.109) that

$$\int t_g^{0\alpha}\,dS_\alpha=\int \tilde{t}_g^{0\alpha}\,dS_\alpha+O\left(\frac{1}{r}\right). \tag{1.111}$$

Thus, the equivalence of the canonical and symmetric energy-momentum tensors for computing the integral characteristics of gravitational radiation is evident from expressions (1.110) and (1.111).

As will be shown in Sec. 1.10, the components $\tilde{t}_g^{00}$ and $\tilde{t}_g^{0\alpha}$ are positively signed quantities and, moreover, only the transverse components of the gravitational wave contribute to energy-momentum. As a result, the energy of the source falls due to expression (1.98).

To obtain the symmetric energy-momentum tensor density of matter t_{M}^{ni} in flat space-time, we note that the metric tensor γ_{ni} enters the Lagrangian density of matter only via the metric tensor of the Riemannian space-time. Hence, the tensor density t_{M}^{ni} may be written as

$$\begin{aligned} t_{\mathrm{M}}^{ni} &= T^{ni}\,[1 - f/2 + (b_3/4)\, f_{lm} f^{lm} + (b_4/2)\, f^2] + f^{ni} T^{lm} \gamma_{lm} \\ &- [b_1 T^{lm} f_l^n f_m^i + b_2 T^{lm} f_{lm} f^{ni} + 2b_3 f_p^i f^{np} T^{lm} \gamma_{lm} \\ &\qquad + 2b_4 f^{ni} f T^{lm} \gamma_{lm}]/4 - 2\Lambda^{(ni)}. \qquad (1.112) \end{aligned}$$

The expression for Λ_{ni} can be obtained from formula (1.104) provided that we take

$$\begin{aligned} A^{lm} &= -T^{lm}/2 + \gamma^{lm} T^{ni} \gamma_{ni}/4 \\ &- [b_1 T^{ln} f_n^m + b_1 T^{nm} f_n^l + b_2 \gamma^{lm} T^{ni} f_{ni} + b_2 T^{lm} f \\ &\qquad + 2b_3 f^{lm} T^{ni} \gamma_{ni} + 2b_4 \gamma^{lm} f T^{ni} \gamma_{ni}]/8. \qquad (1.113) \end{aligned}$$

1.10 EMISSION OF GRAVITATIONAL WAVES

The field equations which incorporate the accepted gauge will have the form

$$\Box f^{lm} = -16\pi I^{lm} \qquad (1.114)$$

and the tensor current I^{lm} (1.85) is only defined in matter.

Since the metric tensor g_{nm} and the energy-momentum tensor of the free gravitational field (i.e. the field outside the matter) depend only on the fields f_{nm}, field equation (1.114) can be solved for f_{nm}. Suppose we write the tensors f^{lm} and I^{lm} in the form of time-dependent Fourier integrals and we isolate a static part $J_0(\mathbf{r})$ from the spectrum $\tilde{I}^{ln}(\omega, \mathbf{r})$. It is evident that the static part of the tensor current $J_0(\mathbf{r})$ will only yield static solutions, and for this reason we can drop it. The following equations for a gravitational field will thus be obtained in terms of Fourier amplitudes:

$$\Delta \tilde{f}^{nm} + \omega^2 \tilde{f}^{nm} = 16\pi \tilde{I}^{nm}. \qquad (1.115)$$

If the origin of the Cartesian coordinate system is fixed at a certain point in the source, a solution of the field equations may be written down in this coordinate system as

$$\begin{aligned} \tilde{f}^{nm} &= -4 \int \frac{\exp(i\omega R)}{R} \tilde{I}^{nm}(\mathbf{r}', \omega)\, d^3\mathbf{r}', \\ R &= |\mathbf{r} - \mathbf{r}'|. \qquad (1.116) \end{aligned}$$

Making use of Lorentz conditions (1.92), $i\omega\tilde{f}^{0m} = \partial_\alpha f^{\alpha m}$, we can express components $\tilde{f}^{0n}$ in terms of the spatial components:

$$\tilde{f}^{00} = -\frac{1}{\omega^2}\partial_\alpha\partial_\beta\tilde{f}^{\alpha\beta},$$

$$\tilde{f}^{0\alpha} = -\frac{i}{\omega}\partial_\beta\tilde{f}^{\alpha\beta}.$$

Outside the gravitational wave source and by choosing the gauge

$$f^{nm} \to f^{nm} + \partial^n a^m + \partial^m a^n - \gamma^{nm}\partial_l a^l, \tag{1.117}$$

which is consistent with Lorentz condition (1.92) at $\Box a^n = 0$, we may impose four additional conditions on the wave components f'^{nm} for each of the independent gauge vectors. We can choose the following conditions as these $\tilde{f}' = 0$, $\tilde{f}'^{0\alpha} = 0$ (the TT gauge).

As the result of this gauging, we obtain

$$\tilde{f}'^{\alpha\beta} = \tilde{f}^{\alpha\beta} - \gamma^{\alpha\beta}\tilde{f}/2 - \left(\frac{i}{\omega}\right)[\partial^\beta\tilde{f}^{0\alpha} + \partial^\alpha\tilde{f}^{0\beta}] - \frac{1}{\omega^2}\partial^\alpha\partial^\beta[\tilde{f}^{00} - \tilde{f}/2]. \tag{1.118}$$

Allowing for the Lorentz conditions (1.92), these expressions can be written as

$$\tilde{f}'^{\alpha\beta} = \tilde{P}^{\alpha\beta} - \frac{1}{\omega^2}[\partial^\beta\partial_\eta\tilde{P}^{\alpha\eta} + \partial^\alpha\partial_\eta\tilde{P}^{\beta\eta}] + \frac{1}{2\omega^2}\gamma^{\alpha\beta}\partial_\eta\partial_\nu\tilde{P}^{\eta\nu} + \frac{1}{2\omega^4}\partial^\alpha\partial^\beta\partial_\nu\partial_\eta\tilde{P}^{\nu\eta}, \tag{1.119}$$

with the following notation

$$\tilde{P}^{\alpha\beta} = \tilde{f}^{\alpha\beta} - \frac{1}{3}\gamma^{\alpha\beta}\tilde{f}^{\eta}_{\eta}. \tag{1.120}$$

Thus, in the general case, the wave solution of the field equations contains six nonzero spatial components $\tilde{f}'^{\alpha\beta}$ but, by virtue of the three Lorentz conditions (1.92) and the trace $\tilde{f}' = 0$, only two components are independent. These additional conditions are the well known ones for the irreducible spin 2 representation under TT gauge, hence, a free gravitational wave has spin 2 but the scalar component that corresponds to the irreducible representation of spin 0 is not emitted in the form of gravitational waves.

Wave solutions of the gravitational field equations are generally expressed in a somewhat different form that vividly demonstrates the quadrupole character of the emitted gravitational waves.

In our case the solution obtained may also be expressed in terms of the generalized quadrupole moments of the tensor current I^{nm}. To do so, we take into account that the spatial coordinates $\tilde{f}^{\alpha\beta}$

(1.116) may, because the tensor current is conserved ($\partial_n I^{nm} = 0$), be written as

$$\tilde{f}^{\alpha\beta} = 2\omega^2 \left\{ \int \frac{\exp(i\omega R)}{R} \tilde{I}^{00} x^\alpha x^\beta \, dV \right.$$

$$+ \frac{2i}{\omega} \partial_\eta \int \frac{\exp(i\omega R)}{R} \tilde{I}^{0\eta} x^\alpha x^\beta \, dV$$

$$\left. - \frac{1}{\omega^2} \partial_\eta \partial_\gamma \int \frac{\exp(i\omega R)}{R} \tilde{I}^{\eta\gamma} x^\alpha x^\beta \, dV \right\}. \quad (1.121)$$

This relation is rigorous but can be simplified provided the linear dimensions of the source are substantially less than the distance from its center to the point of observation. Neglecting the non-wave terms that diminish faster than $1/r$, we obtain

$$\tilde{f}^{\alpha\beta} = \frac{2\omega^2}{r} \int dV \, x^\alpha x^\beta \exp(i\omega R) [\tilde{I}^{00} + 2n_\varepsilon \tilde{I}^{0\varepsilon} + n_\varepsilon n_\gamma \tilde{I}^{\varepsilon\gamma}],$$

where $n_\alpha = \frac{x_\alpha}{r}$, $n_\alpha n^\alpha = -1$. Expression (1.120) may then be written as

$$\tilde{P}^{\alpha\beta} = \frac{2\omega^2}{r} \int dV \left[x^\alpha x^\beta - \frac{1}{3} \gamma^{\alpha\beta} x_\varepsilon x^\varepsilon \right] \exp(i\omega R) [\tilde{I}^{00} + 2n_\varepsilon \tilde{I}^{0\varepsilon}$$

$$+ n_\varepsilon n_\gamma \tilde{I}^{\varepsilon\gamma}]. \quad (1.122)$$

Introducing the projection operators

$$Z^{\alpha\beta} = \gamma^{\alpha\beta} + n^\alpha n^\beta, \quad (1.123)$$

which satisfy the conditions

$$Z^{\alpha\beta} \gamma_{\alpha\beta} = 2, \quad Z^{\alpha\beta} Z_{\beta\varepsilon} = Z^\alpha_\varepsilon,$$

relation (1.119) can be rewritten as

$$\tilde{f}'^{\alpha\beta} = [Z^\alpha_\varepsilon Z^\beta_\gamma - Z^{\alpha\beta} Z_{\varepsilon\gamma}/2] \tilde{P}^{\varepsilon\gamma}. \quad (1.124)$$

Substituting expression (1.122) into the Fourier integral, we obtain

$$P^{\alpha\beta} = -\frac{2}{r} \frac{d^2}{dt^2} \int dV \left(x^\alpha x^\beta - \frac{1}{3} \gamma^{\alpha\beta} x_\varepsilon x^\varepsilon \right) [I^{00} + 2n_\varepsilon I^{0\varepsilon}$$

$$+ n_\varepsilon n_\gamma I^{\varepsilon\gamma}]_{\text{ret}}. \quad (1.125)$$

Here $[.\ .\ .\ .\ .]_{\text{ret}}$ stands for the bracketed expression being taken at the retarding instant $t' = t - R$. If we introduce traceless tensor of the generalized quadrupole moment

$$\mathscr{D}^{\alpha\beta} = D^{\alpha\beta} + 2n_\varepsilon D^{\alpha\beta\varepsilon} + n_\varepsilon n_\gamma D^{\alpha\beta\varepsilon\gamma}, \quad (1.126)$$

where

$$\left.\begin{aligned} D^{\alpha\beta} &= \int dV\,(3x^{\alpha}x^{\beta} - \gamma^{\alpha\beta}x_{\varepsilon}x^{\varepsilon})\,[I^{00}]_{\text{ret}}, \\ D^{\alpha\beta\varepsilon} &= \int dV\,(3x^{\alpha}x^{\beta} - \gamma^{\alpha\beta}x_{\delta}x^{\delta})\,[I^{0\varepsilon}]_{\text{ret}}, \\ D^{\alpha\beta\varepsilon\gamma} &= \int dV\,(3x^{\alpha}x^{\beta} - \gamma^{\alpha\beta}x_{\delta}x^{\delta})\,[I^{\varepsilon\gamma}]_{\text{ret}}, \end{aligned}\right\} \tag{1.127}$$

the gravitational wave components (1.124) may then be written in the form

$$f'^{\alpha\beta} = -\frac{2}{3r}\left(Z^{\alpha}_{\varepsilon}Z^{\beta}_{\gamma} - \frac{1}{2}Z^{\alpha\beta}Z_{\varepsilon\gamma}\right)\ddot{\mathscr{D}}^{\varepsilon\gamma}. \tag{1.128}$$

Here and from now on the dot denotes differentiation over time.

Since $\partial_{\varepsilon}f_{\alpha\beta} = n_{\varepsilon}\dot{f}_{\alpha\beta}$, the following expression can be obtained for the components of the energy-momentum $\tilde{t}^{\alpha}_{g0}$, $\tilde{t}^{0}_{g0}$ of a gravitational wave:

$$\tilde{t}^{\alpha}_{g0} = n^{\alpha}\tilde{t}^{0}_{g0} = \frac{1}{32\pi}\,n^{\alpha}\dot{f}_{\beta\varepsilon}\dot{f}^{\beta\varepsilon}.$$

Then for the radiation intensity of gravitational wave energy through solid angle element $d\Omega$ we have

$$\frac{dI}{d\Omega} = \frac{1}{32\pi}\,r^{2}\dot{f}_{\alpha\beta}\dot{f}^{\alpha\beta} \geqslant 0. \tag{1.129}$$

This expression shows that the radiation intensity of gravitational wave energy through a solid angle element is a positive quantity for all values of the tensor components $\dot{f}_{\alpha\beta}$, provided none of the components vanish. If all the components $\dot{f}_{\alpha\beta}$ equal zero, then $dI/d\Omega$ too will equal zero.

Making use of relations (1.123) and (1.128), expression (1.129) may be written as

$$\frac{dI}{d\Omega} = \frac{1}{36\pi}\left\{\frac{1}{4}(\dddot{\mathscr{D}}^{\alpha\beta}n_{\alpha}n_{\beta})^{2} + \frac{1}{2}\dddot{\mathscr{D}}_{\alpha\beta}\dddot{\mathscr{D}}^{\alpha\beta} + \dddot{\mathscr{D}}_{\alpha\beta}\dddot{\mathscr{D}}^{\beta\gamma}n_{\gamma}n^{\alpha}\right\}. \tag{1.130}$$

Consider further the case of emission of weak gravitational waves as being the most frequent in practice. In the generally analysed linear approximation the tensor current I^{nm} (1.85) has to be taken with the gravitational field absent. It follows from expressions (1.74) and (1.95) that here

$$I^{nm} = T^{nm}. \tag{1.131}$$

When the emitted gravitational waves have wevelengths substantially exceeding the source's dimensions, retardation in the system may

be neglected and the bracketed expressions in formulas (1.127) may be taken for the instant $t' = t - R$.

The following expression will then be obtained for the energy loss in all directions per unit time:

$$I = -\frac{dE}{dt} = \frac{G}{45c^5} \dddot{D}_{\alpha\beta} \dddot{D}^{\alpha\beta}, \tag{1.132}$$

where

$$D^{\alpha\beta} = \int dV \left(3x^\alpha x^\beta - \gamma^{\alpha\beta} x_\varepsilon x^\varepsilon\right) T^{00}(t - R, \mathbf{r}')$$

the gravitational constant, G, and the velocity of light, c, being introduced explicitly.

This formula agrees with the results [82, 83] of the indirect measurements of the energy loss by the double pulsar system PSR 1913 + 16 assumed to be due to emission of gravitational waves. Since the standard GTR computation of the "energy loss" in terms of energy-momentum pseudotensors using a weak field approximation resulted in expression (1.132), it was concluded [83, 127] that the results of observations agreed with the predictions of Einstein's theory.

However, as has been shown in [5, 9], formula (1.132) was not inferred from Einstein GTR, as in Einstein's theory it is only possible to speak about curvature waves which are the ones associated with the energy transfer of matter. However a common form of conservation law is absent for this case, so it is impossible to calculate the energy lost by the source as well as determining any gravitational wave energy flux using the GTR.

Let us clarify why a GTR calculation of "energy fluxes" of gravitational waves results in physically acceptable formulas for a certain narrow class of the coordinate frames. The main reason is that the theory of gravitation can be constructed in pseudo-Euclidean space-time, and hence possesses energy-momentum conservation laws for matter and a gravitational field when taken in conjunction:

$$D_i \left[t_{\mathbf{g}}^{ik} + t_{\mathbf{M}}^{ik}\right] = 0. \tag{1.133}$$

It is only from this theory that expression (1.129) for the intensity of gravitational radiation may be obtained from the conservation law (1.133) written in Cartesian coordinates:

$$\partial_i \left[t_{\mathbf{g}}^{ik} + t_{\mathbf{M}}^{ik}\right] = 0. \tag{1.134}$$

GTR energy-based calculations usually employ the relation

$$\partial_i \left[-g\left(T^{ik} + \tau^{ik}\right)\right] = 0, \tag{1.135}$$

where τ^{ik} is the Landau-Lifshitz energy-momentum pseudotensor:

$$\tau^{ik}=\frac{c^4}{16\pi G}\{(2\Gamma^n_{ml}\Gamma^p_{np}-\Gamma^n_{lp}\Gamma^p_{mn}-\Gamma^n_{nl}\Gamma^p_{mp})(g^{il}g^{km}-g^{ik}g^{lm})$$
$$+g^{il}g^{mn}(\Gamma^k_{pl}\Gamma^p_{mn}+\Gamma^k_{mn}\Gamma^p_{pl}-\Gamma^k_{np}\Gamma^p_{ml}-\Gamma^k_{ml}\Gamma^p_{np})$$
$$+g^{kl}g^{mn}(\Gamma^i_{pl}\Gamma^p_{mn}+\Gamma^i_{mn}\Gamma^p_{pl}-\Gamma^i_{np}\Gamma^p_{ml}-\Gamma^i_{ml}\Gamma^p_{np})$$
$$+g^{lm}g^{np}(\Gamma^i_{nl}\Gamma^k_{mp}-\Gamma^i_{ml}\Gamma^k_{np})\}. \qquad (1.136)$$

This naturally differs from the covariant conservation law (1.133) of the field theory of gravitation. Yet expressions (1.134) and (1.135) have the same form in Cartesian coordinates. Moreover, using the lowest non-vanishing approximation, expression (1.136) for the components $\tau^{0\alpha}$ of the energy-momentum pseudotensor, written in Cartesian coordinates, coincides with the expression for the components of the energy-momentum tensor from the field theory of gravitation:

$$(-g)\,\tau^{0\alpha}=t_g^{0\alpha}=\frac{G\dot{h}_{\beta\varepsilon}\dot{h}^{\beta\varepsilon}}{32\pi c^6r^2}\,n^\alpha+O\left(\frac{1}{r^3}\right).$$

Consequently, it is as if the GTR and the field theory of gravitation are connected in the Cartesian coordinate system. This is why we obtain expression (1.30) for the "intensity of gravitational radiation" and (1.31) for the "total intensity" from the GTR in a narrow class of the coordinate systems that are close to the Cartesian system.

This circumstance produces an illusion that formulas (1.30) and (1.31) for calculating the energy loss are inferred from the GTR. However, because the transformation law for the energy-momentum tensor obtained from the field theory of gravitation differs from that for energy-momentum pseudotensor obtained from the GTR in other coordinate frames their expressions differ even in the lowest nonvanishing approximation. As has been shown in [9], calculation of the "intensity of gravitational radiation" using expression (1.135) leads to physically absurd results. However, energy fluxes of gravitational radiation using the field theory of gravitation has, by virtue of the tensor character of the conservation laws (1.133), a rigorous, definite physical meaning in any admissible frame of reference.

For a transformation from Cartesian coordinates x^α_{old} to new ones x^α_{new}, the new coordinates would be related to the old by the expression

$$x^\alpha_{\text{old}}=x^\alpha_{\text{new}}\left[1+\frac{G}{c^5}\frac{F(ct_{\text{new}}-r_{\text{new}}\theta_{\text{new}},\varphi_{\text{new}})}{r^{3/2}_{\text{new}}}\right.$$
$$\left.\times(1-\exp(-\varepsilon^2r^2_{\text{new}}))\right], \qquad (1.137)$$

$$t_{\text{old}}=t_{\text{new}},$$

where F is some arbitrary function bounded for every value of $u = ct_{\text{new}} - r_{\text{new}}$ and the angles θ_{new} and φ_{new} such that $\max |F| = A < \infty$.

It can readily be checked that transformation (1.137) corresponds to a change in the arithmetization of the points in a three-dimensional space along the radius:

$$r_{\text{old}} = r_{\text{new}} \left[1 + \frac{G}{c^5} \frac{F(u, \theta_{\text{new}}, \varphi_{\text{new}})}{r_{\text{new}}^{3/2}} (1 - \exp(-\varepsilon^2 r_{\text{new}}^2)) \right]. \tag{1.138}$$

To do this, set down the identities inferred from expressions (1.137) and (1.138):

$$\cos\theta_{\text{new}} = \frac{z_{\text{new}}}{r_{\text{new}}} = \frac{z_{\text{old}}}{r_{\text{old}}} = \cos\theta_{\text{old}};$$

$$\sin\theta_{\text{new}} \cos\varphi_{\text{new}} = \frac{x_{\text{new}}}{r_{\text{new}}} = \frac{x_{\text{old}}}{r_{\text{old}}} = \sin\theta_{\text{old}} \cos\varphi_{\text{old}};$$

$$\sin\theta_{\text{new}} \sin\varphi_{\text{new}} = \frac{y_{\text{new}}}{r_{\text{new}}} = \frac{y_{\text{old}}}{r_{\text{old}}} = \sin\theta_{\text{old}} \sin\varphi_{\text{old}}.$$

Whence we obtain

$$\theta_{\text{new}} = \theta_{\text{old}}, \quad \varphi_{\text{new}} = \varphi_{\text{old}}.$$

Consequently, the values of the angles θ and φ at any spatial point in the new frame coincide with those in the old one. It should be stressed too highly that Cartesian coordinates are the basic ones for us, spherical coordinates r, θ and φ being considered no more than convenient notations for certain combinations of Cartesian coordinates.

Since the function F is time-dependent, transformation (1.137) describes, for the general case, the transition to frame of reference that is moving radially relative to the old frame.

For a transformation to have an inverse and to be one-to-one, it is necessary and sufficient that the following condition is satisfied:

$$\frac{\partial r_{\text{old}}}{\partial r_{\text{new}}} > 0. \tag{1.139}$$

The transformation Jacobian (1.137), too, will then not vanish:

$$J = \det \left\| \frac{\partial x_{\text{old}}}{\partial x_{\text{new}}} \right\| = \left[1 + \frac{GF}{c^5 r_{\text{new}}^{3/2}} (1 - \exp(-\varepsilon^2 r_{\text{new}}^2)) \right]^2 \times \frac{\partial r_{\text{old}}}{\partial r_{\text{new}}} > 0.$$

It can easily be verified that by choosing an appropriate value of ε, the condition can be satisfied.

Now when the components $\tau^{0\alpha}$ of the energy-momentum pseudotensor (1.136) are calculated using the new coordinates, we obtain

$$-g\tau^{0\alpha}=\frac{G}{32\pi c^6 r^2}n^\alpha\left[\dot{h}_{\beta\varepsilon}\dot{h}^{\beta\varepsilon}-8\frac{\partial F}{\partial u}\frac{\partial^2 F}{\partial u^2}\right]+O\left(\frac{1}{r^{5/2}}\right). \tag{1.140}$$

As a consequence, the "intensity of gravitational radiation" in a solid angle element will in the new coordinates have the form

$$\frac{dI}{d\Omega}=\frac{G}{32\pi c^5}\left[\dot{h}_{\beta\varepsilon}\dot{h}^{\beta\varepsilon}-8\frac{\partial F}{\partial u}\frac{\partial^2 F}{\partial u^2}\right].$$

Using the relation

$$h^{\alpha\beta}=-\frac{2}{3}\left[Z_\varepsilon^\alpha Z_\gamma^\beta-\frac{1}{2}Z^{\alpha\beta}Z_{\varepsilon\gamma}\right]\ddot{D}^{\varepsilon\gamma},$$

we get

$$\frac{dI}{d\Omega}=\frac{G}{36\pi c^5}\left\{\frac{1}{4}(\dddot{D}_{\alpha\beta}n^\alpha n^\beta)^2+\frac{1}{2}\dddot{D}^{\alpha\beta}\dddot{D}_{\alpha\beta}\right.$$
$$\left.+\dddot{D}_{\alpha\beta}\dddot{D}^{\beta\gamma}n^\alpha n_\gamma-9\frac{\partial F}{\partial u}\frac{\partial^2 F}{\partial u^2}\right\}. \tag{1.141}$$

Integrating this expression over all directions, we obtain the "total radiation" in the new coordinates:

$$I=\frac{G}{45c^5}\left[\dddot{D}_{\alpha\beta}\dddot{D}^{\alpha\beta}-\frac{45}{4\pi}\int d\Omega\frac{\partial F}{\partial u}\frac{\partial^2 F}{\partial u^2}\right]. \tag{1.142}$$

Thus, both the intensity of gravitational radiation and the total intensity calculated from the GTR using the energy-momentum pseudotensor depend on the choice of coordinates and even when the metric is asymptotically Galilean they may be reduced to a physically absurd form by an appropriate choice of coordinates.

Particularly in the aforementioned example, quantities (1.141) and (1.142) can be made to vanish or become negative over the whole region of space confined in the shell between two spheres of radii $r_1=ct-u_1$ and $r_2=ct-u_2$. On the other hand, the intensity of gravitational radiation through any element of a spherical surface having an arbitrary radius r, as well as the total intensity of gravitational radiation through the sphere, can be made both to vanish and go negative within any prescribed time interval

$$t_1=\frac{r+u_1}{c}<t<\frac{r+u_2}{c}=t_2.$$

Yet the components $\tilde{t}_g^{0\alpha}$ of the energy-momentum tensor in the field theory of gravitation will, when transformed to the new coordinates (1.137), have the form

$$\tilde{t}_g^{0\alpha}=\frac{Gn^\alpha}{32\pi c^6 r^2}\dot{h}_{\beta\varepsilon}\dot{h}^{\beta\varepsilon}\left[1+O\left(\frac{1}{\sqrt{r}}\right)\right]. \tag{1.143}$$

Consequently, expressions for the intensity of gravitational radiation and the total intensity obtained from the field theory of gravitation in the new frame will be the same as the corresponding expressions (1.130) and (1.132) obtained in the old frame:

$$\frac{dI}{d\Omega}=\frac{G}{32\pi c^5}\dot{h}_{\beta\varepsilon}\dot{h}^{\beta\varepsilon},$$
$$I=\frac{G}{45c^5}\dddot{D}_{\alpha\beta}\dddot{D}^{\alpha\beta}. \tag{1.144}$$

Note also that, in both frames, the asymptotic expression for the curvature tensor components as $r\to\infty$ will be the same:

$$\left.\begin{aligned}
R_{0\alpha 0\beta} &= -\frac{G}{2c^6 r}\ddot{h}_{\alpha\beta};\\
R_{\alpha\beta 0\varepsilon} &= \frac{G}{2c^6 r}\left(\ddot{h}_{\alpha\varepsilon}n_\beta-\ddot{h}_{\beta\varepsilon}n_\alpha\right);\\
R_{\alpha\beta\gamma\varepsilon} &= \frac{G}{2c^6 r}\Big(\ddot{h}_{\alpha\varepsilon}n_\beta n_\gamma+\ddot{h}_{\beta\gamma}n_\alpha n_\varepsilon-\ddot{h}_{\alpha\gamma}n_\beta n_\varepsilon\\
&\qquad\qquad -\ddot{h}_{\beta\varepsilon}n_\alpha n_\gamma\Big).
\end{aligned}\right\} \tag{1.145}$$

Thus, the Einstein's quadrupole formula for the intensity of gravitational radiation does not result from the GTR, since, as we saw, the "energy lost" by gravitational radiation can vanish or even become negative, depending on the choice of reference frame. As a consequence, this formula cannot be used for any energy-based computations in the GTR. Thus, Einstein's theory cannot indicate the cause of the observed [83] energy loss by the double pulsar system PSR 1913 + 16.

A gravitational field in the field theory of gravitation possesses energy-momentum like other physical fields and, when weak gravitational waves are emitted by a slowly moving source, the energy of the source is decreased according to formula (1.132). For this reason experimental proof that a gravitational field is a physical one that transmits energy, thus decreasing a source energy, would confirm the notions developed here, since the GTR cannot explain an energy loss by matter by gravitational wave emission.

Concluding this section, we will briefly discuss the problem of computing the Riemann tensor in the field theory of gravitation. It is possible in Einstein's theory [2, 4] for the energy-momentum pseudotensor of gravitational waves to vanish without the components of the Riemann tensor vanishing. This fact was due witness to the fallacy of interpreting energy-momentum pseudotensors as energy characteristics of a gravitational field.

In the field theory of gravitation the Riemann tensor is always equal to zero, provided the components of the energy-momentum tensor of gravitational waves vanish, i.e. the Riemannian space-

time always needs the energy and momentum of a gravitational field to form. Note that the Riemannian space-time metric is only meaningful inside the matter. The components g_{ni} of the metric tensor and the curvature tensor R^i_{nlm} can be calculated at any point, including those outside the matter. However the necessity for gauging the field in a due manner outside the matter must be allowed for since physical quantities are independent of the field components f_{nm} which are changing under a gauge transformation. These components do not enter the expression for the gravitational field energy-momentum tensor. They always can be made to vanish by an appropriate gauge transformation. Hence, while geometric characteristics of space-time outside matter (e.g., of the metric tensor g_{ni} and of the Riemann tensor R^i_{nlm}) are calculated, we have to put only those components f_{nm} which enter the energy-momentum tensor of gravitational field into the coupling equation (1.95). It will be assumed that all other field components vanish, since they can be made to do so with an appropriate gauge transformation. Thus, our theory will always be intrinsically self-consistent.

Let all components of the canonical energy-momentum tensor of the free gravitational waves vanish. Then, from expression (1.101) at $n = p = 0$ we obtain

$$\dot{f}_{lm}\dot{f}^{lm} - \frac{1}{2}\dot{f}^2 = 0, \tag{1.146}$$

We can show that under TT gauge, all the components of free gravitational waves are always zero by condition (1.146). All components of the Riemannian space-time metric tensor under this gauging coincide then with the components of the flat space-time metric tensor, i.e. $g_{ni} = \gamma_{ni}$. For this reason, the curvature tensor also vanishes when the energy-momentum tensor of a gravitational field vanish.

Consider a certain point. Suppose the Cartesian axis x goes through the point of observation and we fix a sufficiently small region about this point so that the gravitational wave can be taken as flat within this region. Every wave component will then depend only on the difference $t - x$. In this case the conditions $\partial_n f^{nm} = 0$ will take the form

$$\dot{f}^{00} = \dot{f}^{01} = \dot{f}^{11}, \quad \dot{f}^{02} = \dot{f}^{12}, \quad \dot{f}^{03} = \dot{f}^{13}.$$

Integrating these equations and assuming the integration constants do not vanish, since gravitational waves do not have a time-dependent part, we obtain

$$f^{00} = f^{01} = f^{11}, \quad f^{02} = f^{12}, \quad f^{03} = f^{13}.$$

All these components vanish by TT gauge. Besides, since the trace vanishes, $f^n_n = 0$, we have $f^{22} = -f^{33}$.

From the condition that the energy-momentum tensor vanishes (1.101), we obtain

$$2(\dot{f}_{23})^2 + \frac{1}{2}(\dot{f}_{22} - \dot{f}_{33})^2 = 0.$$

The transverse components of gravitational wave will then vanish as well:

$$f_{23} = f_{22} = f_{33} = 0.$$

Thus, under TT gauge from the condition the energy-momentum tensor of a free gravitational wave must vanish we find all the components of this wave are zero. Hence, all the components of the Riemannian space-time metric tensor too will coincide with the corresponding components of the pseudo-Euclidean space-time metric tensor, i.e. $g_{ni} = \gamma_{ni}$. This means that all the components of the Riemannian tensor will vanish, $R^i_{mln} = 0$.

1.11 A POST-NEWTONIAN APPROXIMATION OF THE FIELD THEORY OF GRAVITATION

To make a comparison between the results of the experiments performed within the solar system and the predictions of the different metric theories of gravitation easier, Nordtvedt and Will [84] have developed a formalism called parametrised post-Newtonian (PPN).

Using this formalism, the Riemannian space-time metric produced by a body of a perfect liquid is written using every possible generalized gravitational potential having arbitrary coefficients called post-Newtonian parameters. By using the revised Will-Nordtvedt parameters, the Riemannian space-time metric may be written in the form

$$\left.\begin{aligned}
g_{00} &= 1 - 2U + 2\beta U^2 - (2\gamma + 2 + \alpha_3 + \xi_1)\Phi_1 + \xi_1 A \\
&+ 2\xi_w \Phi_w - 2\left[(3\gamma + 1 - 2\beta + \xi_2)\Phi_2 \right. \\
&\left. + (1 + \xi_3)\Phi_3 + 3(\gamma + \xi_4)\Phi_4\right] \\
&- (\alpha_1 - \alpha_2 - \alpha_3) w_\alpha w^\alpha U + \alpha_2 w^\alpha w^\beta U_{\alpha\beta} \\
&- (2\alpha_3 - \alpha_1) w^\alpha V_\alpha; \\
g_{0\alpha} &= \frac{1}{2}(4\gamma + 3 + \alpha_1 - \alpha_2 + \xi_1) V_\alpha \\
&+ \frac{1}{2}(1 + \alpha_2 - \xi_1) W_\alpha - \frac{1}{2}(\alpha_1 - 2\alpha_2) w_\alpha U \\
&+ \alpha_2 w^\beta U_{\alpha\beta}; \\
g_{\alpha\beta} &= (1 + 2\gamma U)\gamma_{\alpha\beta},
\end{aligned}\right\} \quad (1.147)$$

where w^α are the spatial components of the frame velocity with respect to some universal rest frame. For some theories of gravitation this is the center-of-mass velocity of the solar system with respect to the universal rest frame.

The generalized gravitational potentials have the form

$$\left.\begin{aligned}
&U(\mathbf{r},\ t)=\int \frac{\rho_0(\mathbf{r}',\ t)}{R}\,dV, \qquad R=|\mathbf{r}-\mathbf{r}'|,\\
&\Phi_1=-\int \frac{\rho_0 v_\alpha v^\alpha}{R}\,dV, \quad \Phi_2=\int \frac{\rho_0 U}{R}\,dV,\\
&\Phi_3=\int \frac{\rho_0 \Pi}{R}\,dV, \qquad \Phi_4=\int \frac{p}{R}\,dV,\\
&A=\int \frac{\rho_0 v_\alpha v_\beta R^\alpha R^\beta}{R^3}\,dV, \quad V_\alpha=-\int \frac{\rho_0 v_\alpha}{R}\,dV,\\
&W_\alpha=\int \frac{\rho_0 v_\beta R^\beta R_\alpha}{R^3}\,dV, \quad U_{\alpha\beta}=\int \frac{\rho_0 R_\alpha R_\beta}{R^3}\,dV,\\
&\Phi_w=\int \frac{\rho_0(\mathbf{r}',\ t)\,\rho_0(\mathbf{r}'',\ t)}{|\mathbf{r}-\mathbf{r}'|^3}\,(x^\alpha-x'^\alpha)\\
&\qquad\times\left[\frac{x_\alpha-x''_\alpha}{|\mathbf{r}'-\mathbf{r}''|}-\frac{x'_\alpha-x''_\alpha}{|\mathbf{r}-\mathbf{r}''|}\right]d^3\mathbf{r}'\,d^3\mathbf{r}'',\\
&R^\alpha=x^\alpha-x'^\alpha,
\end{aligned}\right\} \tag{1.148}$$

where ρ_0 is the invariant mass density of the body, v^α the velocity components of the perfect liquid elements, p the isotropic pressure, $\rho_0\Pi$ the internal energy density of the perfect liquid.

Every metric theory of gravitation will have an appropriate set of the values of the ten parameters:

$$\beta,\ \gamma,\ \alpha_1,\ \alpha_2,\ \alpha_3,\ \xi_1,\ \xi_2,\ \xi_3,\ \xi_4,\ \xi_w.$$

Hence, from the point of view of experiments performed within the solar system, one theory of gravitation will differ from another only in the values of these parameters. In order to discover which theories of gravitation can describe, in the post-Newtonian limit, the experiments performed in the solar system, it is sufficient to determine the values of the ten post-Newtonian parameters from these experiments and to select only those theories whose post-Newtonian approximations result in values of the parameters which agree with the experiments. These theories of gravitation will then be indistinguishable from the point of view of any experiments carried out with the post-Newtonian accuracy.

Subsequent selection of a theory of gravitation, which matches reality, will be associated either with an increase in the accuracy of measurements up to the post-Newtonian level or by studying the properties of gravitational wave and phenomena in a strong gravitational field.

We shall define a set of values for the post-Newtonian parameters that correspond to the field theory of gravitation.

The gravitational field equations of this theory that will be used to compute the post-Newtonian approximation will be written as

$$\Box^2 f^{nm} = -16\pi J^{nm}; \quad \Box = \partial_i \partial^i. \tag{1.149}$$

By using the notations (1.74), the tensor current can be expressed as

$$J^{nm} = \Box\, h^{nm} - \partial^n \partial_l h^{lm} - \partial^m \partial_l h^{ln} + \gamma^{nm} \partial_l \partial_i h^{li}. \tag{1.150}$$

Following Fock [13] for constructing a post-Newtonian approximation valid in the solar system, we shall consider an astronomical problem. It will be assumed that the components of the energy-momentum tensor of matter vanish in all but a few regions of space. Inside each of these regions the energy-momentum tensor has to fit the accepted model of a perfect liquid and satisfy the covariant conservation equation in Riemannian space-time. In addition to the physical properties of the model of celestial bodies, the energy-momentum tensor of matter will depend on the Riemannian space-time metric as well. The construction of the energy-momentum of matter and the determination of the Riemannian space-time metric tensor should, therefore, be performed simultaneously.

Let us take advantage of the circumstance that within the solar system the maximum values of the gravitational potential, the square v^2 of a characteristic velocity (the velocity of celestial bodies with respect to the center-of-mass of the solar system), the specific pressure p/ρ_0 and the specific internal energy Π have approximately the same order of smallness, ε^2, where $\varepsilon \sim 10^{-3}$ is some dimensionless parameter. Hence, the following estimates will be valid for the solar system:

$$U = O(\varepsilon^2); \quad v^\alpha = O(\varepsilon); \quad \Pi = O(\varepsilon^2); \quad p/\rho_0 = O(\varepsilon^2). \tag{1.151}$$

Also, we are going to deal with the near-zone field, i.e. at distances from the Sun much shorter than the wavelength of the gravitational wave emitted by objects in the solar system, moving at the characteristic velocity $v \sim \varepsilon$: $R/\lambda \sim R\partial/\partial t \sim \varepsilon$. In this case the changes over time of all quantities are primarily driven by the motion of matter and the partial derivatives with respect to time are therefore small compared to position dependent partial derivatives

$$\partial/\partial t = O(\varepsilon)\, \partial/\partial x^\alpha. \tag{1.152}$$

The problem of how to determine the energy-momentum tensor of matter and the Riemannian space-time metric tensor simultaneously will be solved in consecutive stages, each stage corresponding to an expansion of the exact equations in powers of the dimensionless parameter ε.

We have the following rigorous relationships:
the energy-momentum tensor density of a perfect liquid

$$T^{nm} = \sqrt{-g}\,[(p+\mathscr{E})\,u^n u^m - p g^{nm}]; \tag{1.153}$$

the covariant continuity equation

$$(1/\sqrt{-g})\,\partial/\partial x^i\,[\sqrt{-g}\,\rho_0 u^i] = 0; \tag{1.154}$$

and the conservation equation for the energy-momentum tensor density of matter in Riemannian space-time

$$\nabla_n T^{nm} = \partial_n T^{nm} + \Gamma^m_{nl} T^{nl} = 0, \tag{1.155}$$

where $\mathscr{E}$ is the total energy density of the perfect liquid, and u^i the velocity four-vector.

It is more convenient for the present to write the gravitational field equations (1.149) and the minimum-coupling equations (1.95) as

$$\Box^2 \chi^{nm} = -16\pi A^{nm}, \tag{1.156}$$

$$g_{nm} = \gamma_{nm} + \chi_{nm} + \frac{1}{4}\,[b_1 \chi_{nl}\chi^l_m + b_3 \gamma_{nm}\chi^{li}\chi_{li} - (b_1+b_2)\,\chi_{nm}\chi + (b_4 + b_2/2 + b_1/4)\,\chi^2\gamma_{nm}], \tag{1.157}$$

where

$$\chi_{nm} = f_{nm} - \frac{1}{2}\gamma_{nm} f, \quad \chi = \chi^n_n, \tag{1.158}$$

$$A^{nm} = \Box\left[h^{nm} - \frac{1}{2}\gamma^{nm} h^l_l\right] - \partial^n \partial_l h^{lm} - \partial^m \partial_l h^{ln}. \tag{1.159}$$

Let us expand all the quantities from equations (1.155)-(1.158) into power series in the small parameter ε. By neglecting the energy lost through the emission of gravitational waves, these equations must hold for sign inversion of time as well. In doing this, i.e. by transforming the coordinates $x^{0\prime} = -x^0$, the components v^α, $\chi^{0\alpha}$, $T^{0\alpha}$, $g_{0\alpha}$, $A^{0\alpha}$, $\partial/\partial x^0$ change sign. Since $v \sim \varepsilon$ and $\partial/\partial x^0 \sim \varepsilon\partial/\partial x^\alpha$, then by inverting the sign of time the dimensionless parameter ε changes its sign as well. Whence it follows that, by neglecting the energy lost through the emission of gravitational waves, the expansions of the components v^α, $\chi^{0\alpha}$, $T^{0\alpha}$, $g_{0\alpha}$, $A^{0\alpha}$ only contain odd powers of the parameter ε.

Let the expansions of the field χ^{nm} and of the current tensor A^{nm} be written in the form

$$\chi^{nm} = \overset{(1)}{\chi}{}^{nm} + \overset{(2)}{\chi}{}^{nm} + \ldots, \tag{1.160}$$

$$A^{nm} = \overset{(0)}{A}{}^{nm} + \overset{(1)}{A}{}^{nm} + \ldots, \tag{1.161}$$

where the components of the zeroth, $\overset{(0)}{A}{}^{nm}$, first, $\overset{(1)}{A}{}^{nm}$, and second, $\overset{(2)}{A}{}^{nm}$, approximations have the following orders of smallness:

$$\left.\begin{aligned}
&\overset{(0)}{A}{}^{0\alpha}=O(\varepsilon), && \overset{(0)}{A}{}^{00}=O(1), && \overset{(0)}{A}{}^{\alpha\beta}=O(1),\\
&\overset{(1)}{A}{}^{0\alpha}=O(\varepsilon^3), && \overset{(1)}{A}{}^{\delta\alpha}=O(\varepsilon^2), && \overset{(1)}{A}{}^{00}=O(\varepsilon^2),\\
&\overset{(2)}{A}{}^{0\alpha}=O(\varepsilon^5), && \overset{(2)}{A}{}^{00}=O(\varepsilon^4), && \overset{(2)}{A}{}^{\alpha\beta}=O(\varepsilon^4).
\end{aligned}\right\} \tag{1.162}$$

The gravitational field equations (1.156), taking into account expansions (1.160), (1.161) and estimates (1.152) can be rewritten in the form of a number of successive approximations:

$$\Delta^2\overset{(1)}{\chi}{}^{nm}=-16\pi\overset{(0)}{A}{}^{nm}, \tag{1.163}$$

$$\Delta^2\overset{(2)}{\chi}{}^{nm}=-16\pi\overset{(1)}{A}{}^{nm}+2\frac{\partial^2}{\partial t^2}\Delta\overset{(1)}{\chi}{}^{nm}. \tag{1.164}$$

From expressions (1.74) and (1.95), we get

$$\begin{aligned}
h^{nm}=T^{nm}&+\frac{b_1}{4}\left[T^{nl}\chi^m_l+T^{ml}\chi^n_l\right]-\frac{b_1+b_2}{4}\chi T^{nm}\\
&+\frac{b_3}{2}\chi^{nm}T^{li}\gamma_{li}-\frac{b_1+b_2}{4}\gamma^{nm}T^{li}\chi_{li}\\
&+\left(\frac{b_2}{4}+\frac{b_1}{8}+\frac{b_4}{2}\right)\gamma^{nm}\chi T^{li}\gamma_{li}.
\end{aligned} \tag{1.165}$$

Then for the tensor current A^{nm} we shall have

$$\left.\begin{aligned}
\overset{(0)}{A}{}^{nm}&=-\Delta\left[\overset{(0)}{T}{}^{nm}-\frac{1}{2}\gamma^{nm}\overset{(0)}{T}{}^{li}\gamma_{li}\right],\\
\overset{(1)}{A}{}^{nm}&=\frac{\partial^2}{\partial t^2}\left[\overset{(0)}{T}{}^{nm}-\frac{1}{2}\gamma^{nm}\overset{(0)}{T}{}^{li}\gamma_{li}\right]\\
&+\partial^n(\overset{(1)}{\Gamma}{}^m_{li}\overset{(0)}{T}{}^{li})+\partial^m(\overset{(1)}{\Gamma}{}^n_{li}\overset{(0)}{T}{}^{li})\\
&-\Delta\left[\overset{(1)}{T}{}^{nm}-\frac{1}{2}\gamma^{nm}\overset{(1)}{T}{}^{li}\gamma_{li}\right.\\
&+\frac{b_1}{4}\left(\overset{(0)}{T}{}^{nl}\overset{(1)}{\chi}{}^m_l+\overset{(0)}{T}{}^{ml}\overset{(1)}{\chi}{}^n_l\right)-\frac{b_1+b_2}{4}\overset{(0)}{T}{}^{nm}\overset{(1)}{\chi}\\
&+\frac{b_2}{4}\gamma^{nm}\overset{(0)}{T}{}^{li}\overset{(1)}{\chi}_{li}+\frac{b_3}{2}\overset{(1)}{\chi}{}^{nm}\overset{(0)}{T}{}^{li}\gamma_{li}\\
&\left.-\left(\frac{b_2}{8}+\frac{b_3}{4}+\frac{b_4}{2}\right)\gamma^{nm}\overset{(0)}{T}{}^{li}\gamma_{li}\chi\right],
\end{aligned}\right\} \tag{1.166}$$

where $\Delta=-\partial^\alpha\partial_\alpha$.

To determine the post-Newtonian parameters, it will suffice to find components $g_{\alpha\beta}$ correct to ε^2, $g_{0\alpha}$, correct to ε^3, and g_{00}, correct to ε^4. The coupling equation (1.157) suggests that this requires that the field components $\chi^{\alpha\beta}$ are determined correct to ε^2, $\chi^{0\alpha}$, to ε^3, and χ^{00}, to ε^4.

With the initial approximation the Riemannian space-time metric tensor is assumed to coincide with the metric tensor of pseudo-Euclidean space-time, i.e. the gravitational forces are neglected. Equations (1.154) and (1.155) then take the form

$$\partial/\partial x^i (\rho u^i) = O(\varepsilon^2), \quad \partial_n T^{n0} = O(\varepsilon^3), \quad \partial_n T^{n\alpha} = O(\varepsilon^2). \tag{1.167}$$

By taking into consideration estimates (1.151), we infer from these equations that

$$T^{00} = \rho_0 [1 + O(\varepsilon^2)], \quad T^{\alpha\beta} = \rho_0 O(\varepsilon^2), \quad T^{0\alpha} = \rho_0 v^\alpha [1 + O(\varepsilon^2)].$$

Therefore, the tensor current components A^{nm} may be written in the zeroth approximation as

$$\overset{(0)}{A}{}^{00} = -\Delta\rho_0/2, \quad \overset{(0)}{A}{}^{0\alpha} = -\Delta(\rho_0 v^\alpha), \quad \overset{(0)}{A}{}^{\alpha\beta} = \gamma^{\alpha\beta}\Delta\rho_0/2. \tag{1.168}$$

Then from equation (1.163) we shall obtain

$$\overset{(1)}{\chi}{}^{00} = -2U, \quad \overset{(1)}{\chi}{}^{\alpha\beta} = 2U\gamma^{\alpha\beta}, \quad \overset{(1)}{\chi}{}^{0\alpha} = 4V^\alpha. \tag{1.169}$$

As a result, the Riemannian space-time metric tensor components (1.157) may be written in the first approximation as

$$\left.\begin{aligned} g_{00} &= 1 - 2U + O(\varepsilon^4), \\ g_{0\alpha} &= 4V_\alpha [1 + O(\varepsilon^2)], \\ g_{\alpha\beta} &= \gamma_{\alpha\beta} [(1 + 2U) + O(\varepsilon^4)]. \end{aligned}\right\} \tag{1.170}$$

Knowing the metric, this approximation allows the energy-momen tum tensor components of matter to be determined in the next approximation. Using expressions (1.170), it will be found that

$$\left.\begin{aligned} &\sqrt{-g} = 1 + 2U + O(\varepsilon^4), && u^0 = 1 + U - v_\alpha v^\alpha/2, \\ &\Gamma^0_{00} = -\partial U/\partial t + O(\varepsilon^5), && \Gamma^\alpha_{00} = \partial^\alpha U + O(\varepsilon^4), \\ &\Gamma^0_{0\alpha} = -\partial_\alpha U + O(\varepsilon^4), && \Gamma^\alpha_{\beta\gamma} = O(\varepsilon^2), \\ &\Gamma^\alpha_{0\beta} = O(\varepsilon^3), && \Gamma^0_{\alpha\beta} = O(\varepsilon^2). \end{aligned}\right\} \tag{1.171}$$

Let us also introduce a conserved mass density ρ by the identity $\rho = \sqrt{-g}\rho_0 u^0$. In order to obtain the metric in the next approximation, we must construct the energy-momentum tensor density of

matter which would satisfy the conservation equations (1.155) due to the covariant continuity equation

$$\frac{1}{\sqrt{-g}}\left[\frac{\partial\rho}{\partial t}+\frac{\partial}{\partial x^\alpha}(\rho v^\alpha)\right]=0 \quad (1.172)$$

and to the equation of motion of a perfect liquid in the Newtonian approximation [13]:

$$\rho\frac{dv^\alpha}{dt}=\gamma^{\alpha\beta}\left[-\rho\frac{\partial U}{\partial x^\beta}+\frac{\partial p}{\partial x^\beta}\right]+\rho O(\varepsilon^4),$$

$$\rho\, d\Pi/dt=-p\partial_\alpha v^\alpha, \quad d/dt=\partial/\partial t+v^\beta\partial/\partial x^\beta.$$

We can see that these conditions are satisfied by the following components of the energy-momentum tensor of matter:

$$T^{00}=\rho\left[1-v_\alpha v^\alpha/2+\Pi+U\right]+\rho O(\varepsilon^4),$$

$$T^{0\alpha}=\rho v^\alpha\left[1-v_\beta v^\beta/2+\Pi+U\right]+pv^\alpha+\rho O(\varepsilon^4),$$

$$T^{\alpha\beta}=\rho v^\alpha v^\beta-p\gamma^{\alpha\beta}+\rho O(\varepsilon^4).$$

Thus between the conserved and invariant mass densities, ρ and ρ_0, respectively we have the following relationship:

$$\rho=\rho_0\left[1+3U-\frac{v_\alpha v^\alpha}{2}\right]+\rho O(\varepsilon^4). \quad (1.173)$$

Before we can obtain the post-Newtonian approximation, $\overset{(2)}{\chi}{}^{00}$ remains to be determined. By allowing for expressions (1.166), (1.169), (1.171), equation (1.164) for the component $\overset{(2)}{\chi}{}^{00}$ takes the form

$$\Delta^2\overset{(2)}{\chi}{}^{00}=8\pi\frac{\partial^2}{\partial t^2}\rho_0+16\pi\Delta\left\{\frac{3}{2}p+\rho_0\left[\frac{\Pi}{2}-v_\alpha v^\alpha\right.\right.$$
$$\left.\left.-2(b_1+b_2+b_3+b_4)U\right]\right\}. \quad (1.174)$$

Solving this equation, we obtain

$$\overset{(2)}{\chi}{}^{00}=-\frac{\partial^2}{\partial t^2}\int\rho_0 R\, dV-4\Phi_1-2\Phi_3$$
$$-6\Phi_4+8(b_1+b_2+b_3+b_4)\Phi_2. \quad (1.175)$$

Using expressions (1.157), (1.169), and (1.175), the Riemannian space-time metric appears in the post-Newtonian approximation as

$$g_{00}=1-2U+2\beta U^2-4\Phi_1+4(\beta-2)\Phi_2-2\Phi_3-6\Phi_4$$
$$-\frac{\partial^2}{\partial t^2}\int\rho_0 R\, dV+O(\varepsilon^6);$$
$$g_{0\alpha}=4V_\alpha+O(\varepsilon^5), \quad (1.176)$$
$$g_{\alpha\beta}=\gamma_{\alpha\beta}(1+2U)+O(\varepsilon^4),$$

where

$$\beta = 2\,(b_1 + b_2 + b_3 + b_4).$$

Now to determine the values of the post-Newtonian parameters of our theory, the coordinates must be transformed into the coordinate frame of the post-Newtonian expansion of the metric (1.147). Once the coordinate transformation

$$x'^n = x^n + \xi^n(x) \tag{1.177}$$

is done and it is assumed that

$$\xi^\alpha(x) \sim O(\varepsilon^2), \quad \xi^0(x) \sim O(\varepsilon^3),$$

the metric (1.176) will in the new coordinates have the form

$$\left.\begin{aligned} g'_{00} &= g_{00} - 2\partial_0\xi_0 + O(\varepsilon^6),\\ g'_{0\alpha} &= g_{0\alpha} - \partial_0\xi_\alpha - \partial_\alpha\xi_0 + O(\varepsilon^5),\\ g'_{\alpha\beta} &= g_{\alpha\beta} - \partial_\alpha\xi_\beta - \partial_\beta\xi_\alpha + O(\varepsilon^4). \end{aligned}\right\} \tag{1.178}$$

A frame where the non-diagonal components g_{ni} of the spatial part of the metric tensor vanish ($g_{12} = g_{13} = g_{23} = 0$) is usually chosen as the "canonical" frame and where the component g_{00} additionally does not contain terms of the form

$$\frac{\partial^2}{\partial t^2} \int \rho_0 R\, dV.$$

These requirements enable the four-vector to be determined with the accuracy required. In our case, a transformation to the "canonical" coordinates necessitates choosing the following four-vector ξ^n:

$$\xi^\alpha(x) = 0, \quad \xi^0(x) = -\frac{1}{2}\frac{\partial}{\partial t}\int \rho_0 R\, dV.$$

By making use of continuity equation (1.172), we shall obtain

$$\partial_\alpha\xi_0 = \frac{1}{2}[V_\alpha - W_\alpha].$$

As a result, we have the following expression for the metric tensor of the effective Riemannian space-time:

$$\begin{aligned} g_{00} &= 1 - 2U + 2\beta U^2 - 4\Phi_1 + 4(\beta - 2)\Phi_2 - 2\Phi_3\\ &\quad - 6\Phi_4 + O(\varepsilon^6),\\ g_{0\alpha} &= (7/2)\, V_\alpha + W_\alpha/2 + O(\varepsilon^5),\\ g_{\alpha\beta} &= \gamma_{\alpha\beta}\,[1 + 2U] + O(\varepsilon^4). \end{aligned} \tag{1.179}$$

Thus, the post-Newtonian approximation of the field theory of gravitation leads to the Riemannian space-time metric (1.179) containing the sole unknown constant β.

When the gravitational-field source is a static, spherically symmetric body of radius a, this metric looks like

$$\left.\begin{aligned} g_{00} &= 1 - 2M/r + 2\beta M^2/r^2 + O(M^3/r^3), \\ g_{\alpha\beta} &= \gamma_{\alpha\beta}(1 + 2M/r) + O(M^2/r^2), \quad g_{0\alpha} = 0 \end{aligned}\right\}, \tag{1.180}$$

where M is the gravitational mass of the field source:

$$M = 4\pi \int_0^a \rho_0 \left[1 + \Pi + \frac{3p}{\rho_0} + 2(2-\beta)U\right] r^2\, dr. \tag{1.181}$$

Making use of the Newtonian virial theorem for static bodies

$$3 \int p\, dV = \frac{1}{2} \int \rho_0 U\, dV$$

and also of relation (1.173) between the conserved and invariant mass densities, expression (1.181) is reduced to

$$M = 4\pi \int_0^a \rho \left[1 + \Pi + \left(\frac{3}{2} - 2\beta\right) U\right] r^2\, dr. \tag{1.182}$$

As will be seen below, for the post-Newtonian expressions for the gravitational (1.182) and the inertial (1.197) mass of a static, spherically symmetric body to be identical it is necessary to set $\beta = 1$. The post-Newtonian parameters for the field theory of gravitation will then have the following magnitudes:

$$\begin{aligned} &\gamma = \beta = 1, \quad \alpha_1 = \alpha_2 = \alpha_3 = 0, \\ &\xi_1 = \xi_2 = \xi_3 = \xi_4 = \xi_w = 0. \end{aligned} \tag{1.183}$$

In comparison, the GTR post-Newtonian parameters have the same magnitudes [84].

Note that all the parameters α and ξ vanish in Einstein's theory. For a long time this was assumed to be a feature of this theory alone and considered one of its achievements. However, as we see, in the field theory of gravitation these parameters also vanish. The remaining parameters are equal to unity both in the GTR and in the field theory of gravitation.

As the post-Newtonian parameters of both Einstein's GTR and the field theory of gravitation coincide, these two theories are indistinguishable from the standpoint of experiments performed at a post-Newtonian level of accuracy of measurements in the solar system's gravitational field.

As has been shown in [85], the disappearance of the three α parameters has a definite physical meaning. Any theory of gravitation having $\alpha_1 = \alpha_2 = \alpha_3 = 0$ does not possess the preferred universal rest frame in the post-Newtonian limit. In this case, during a transformation from the universal rest frame to a moving frame, the

Riemannian space-time metric in the post-Newtonian limit is forminvariant and the velocity w^α of the new coordinate frame with respect to the rest frame will not enter the metric in an explicit form. It follows from expression (1.183) that the field theory of gravitation has no universal preferred rest frame.

A definite physical meaning can also be assigned to the linear dependence of the parameters ξ and α. As has been shown in [86], when the following equations are satisfied:

$$\left.\begin{array}{ll} \alpha_1=\xi_3=0, & \alpha_2-\xi_1-2\xi_w=0, \\ \xi_2=\xi_w; & \alpha_3+\xi_1+2\xi_w=0, \\ 3\xi_4+2\xi_w=0; & \xi_1+2\xi_w=0, \end{array}\right\} \tag{1.184}$$

the post-Newtonian equations of motion can define some quantities which are not dependent on time in the post-Newtonian approximation. Yet these quantities may only be treated as the energy-momentum and the angular momentum of a system (i.e. as the integrals of motion) in theories of gravitation, which possess conservation laws for the energy-momentum tensor of matter and gravitational field.

So, for example, in Einstein's theory relations (1.184) are satisfied, but the time independent quantities are not integrals of motion of a system composed of matter and gravitational field as is demonstrated by rigorously analysis (see also Sec. 1.12).

In the field theory of gravitation an isolated system has all ten conservation laws with their conventional meanings in pseudo-Euclidean space-time, this means that there are ten integrals of motion for the system in the post-Newtonian approximation. The field theory of gravitation has in this approximation ten time independent quantities and the fulfilment of relations (1.184) in the field theory of gravitation supports this conclusion.

1.12 THE CONSERVATION LAWS IN THE POST-NEWTONIAN APPROXIMATION OF THE FIELD THEORY OF GRAVITATION

According to the field theory of gravitation, a gravitational field behaves just like other physical fields in pseudo-Euclidean space-time. It possesses energy-momentum and contributes to the complete energy-momentum tensor density of the system. The covariant conservation law of the complete energy-momentum tensor density in pseudo-Euclidean space-time, when written in Cartesian coordinate frame, has a familiar meaning:

$$\partial_i\,[t_g^{ni}+t_M^{ni}]=0, \tag{1.185}$$

where t_g^{ni} is the symmetric energy-momentum tensor density of a gravitational field (1.105) and t_M^{ni}, that of matter (1.112).

Taking advantage of the differential conservation law (1.185), one may obtain the integral conservation law as

$$-\frac{\partial}{\partial t}\int dV\,[t_g^{0n}+t_M^{0n}]=\int dS_\alpha\,[t_g^{n\alpha}+t_M^{n\alpha}].$$

If there is no matter or a gravitational energy flux through the surface enclosing the volume, i.e.

$$\int dS_\alpha\,[t_g^{\alpha n}+t_M^{\alpha n}]=0, \tag{1.186}$$

we arrive at the conservation law for the complete four-momentum of an isolated system:

$$\frac{d}{dt}P^n=0,$$

where

$$P^n=\int dV\,[t_g^{0n}+t_M^{0n}]. \tag{1.187}$$

Due to the symmetry of the complete energy-momentum tensor density, the angular momentum tensor of the system will be conserved as well:

$$\frac{d}{dt}M^{ni}=0,$$

where

$$M^{ni}=\int dV\,\{x^n\,[t_g^{0i}+t_M^{0i}]-x^i\,[t_g^{0n}+t_M^{0n}]\}. \tag{1.188}$$

Since the components are conserved

$$M^{0\alpha}=x^0\int dV\,[t_g^{0\alpha}+t_M^{0\alpha}]-\int dV x^\alpha\,[t_g^{00}+t_M^{00}]$$

the center-of-mass of an isolated system defined by the formula

$$\begin{aligned}X^\alpha&=\int x^\alpha\,[t_g^{00}+t_M^{00}]\,dV\Big/\int dV\,[t_g^{00}+t_M^{00}]\\&=(P^\alpha t-M^{0\alpha})/P^0,\end{aligned} \tag{1.189}$$

undergoes uniform rectilinear motion at a velocity

$$\frac{d}{dt}X^\alpha=\frac{P^\alpha}{P^0}.$$

Thus, determining the four-momentum P^n (1.187) will suffice to describe the motion of an isolated system composed of matter and gravitational field. Note that in any real system gravitational waves can be emitted due to motion of the system's constituents,

the matter's thermal motion, etc. Any real system can exchange matter with other systems both in the form of electromagnetic radiation and in the form of particles, atoms, etc. For this reason, even in the most general case, matter and gravitational energy fluxes cannot be neglected, since a great number of astrophysical processes exist for which these energy fluxes are crucial and it is only by taking them into account that many astrophysical processes can be understood and predicted. Yet at the same time the isolation condition (1.186) is fulfilled to some degree of accuracy when systems have low intensity matter and gravitational energy fluxes and the same degree of accuracy can then be claimed for the conservation of the system' s four-momentum. It is for such systems that the post-Newtonia s formalism can be used. In this case condition (1.186) for the system n isolation holds for the post-Newtonian approximation and we aɹs' able to determine the conserved four-momentum of the system.

We now proceed to discover that the post-Newtonian expression is for the four-momentum of the system in the field theory of gravitation. The complete symmetric energy-momentum tensor density in flat space-time has the form

$$
\begin{aligned}
t^{ni} = t_{\mathrm{g}}^{ni} + t_{\mathrm{M}}^{ni} = \frac{1}{64\pi}\Big\{ & -\gamma^{ni}\Big[\partial_l f_{ms}\partial^l f^{ms} - \frac{1}{2}\partial_l f\partial^l f\Big] \\
& - \partial^n f\partial^i f + 2\partial^n f_{lm}\partial^i f^{lm} - 2f^{im}\Box f^n_m - 2f^{nm}\Box f^i_m + 2f^{ni}\Box f\Big\} \\
& - \frac{1}{32\pi}\partial_l\{f^i_m\partial^n f^{lm} + f^n_m\partial^i f^{lm} - f^{lm}(\partial^i f^n_m + \partial^n f^i_m)\} - 2\Lambda^{(ni)} \\
& + T^{ni}\Big[1 - \frac{1}{2}f + \frac{b_3}{4}f_{lm}f^{lm} + \frac{b_4}{2}f^2\Big] + \frac{1}{2}f^{ni}\gamma_{lm}T^{lm} \\
& - \frac{1}{4}[b_1 T^{lm}f^i_l f^n_m + b_2 f^{ni}T^{lm}f_{lm} + 2b_3 f^n_s f^{is}T^{lm}\gamma_{lm} \\
& + 2b_4 f^{ni}fT^{lm}\gamma_{lm}],
\end{aligned}
\tag{1.190}
$$

where Λ^{ni} is defined by (1.104) with the tensor A^{lm} here having the form

$$
\begin{aligned}
A^{lm} = -\frac{1}{32\pi}\Big\{ & \Box\Big(f^{lm} - \frac{1}{2}\gamma^{lm}f\Big) \\
& + 16\pi\Big(h^{lm} - \frac{1}{2}\gamma^{lm}h^n_n\Big)\Big\}.
\end{aligned}
\tag{1.191}
$$

It follows from (1.190) that the components t^{00} and $t^{0\alpha}$ of the complete symmetric energy-momentum tensor of the system may be determined accurate up to $t^{00} \sim \rho O(\varepsilon^2)$, $t^{0\alpha} \sim \rho O(\varepsilon^3)$ inclusively.

We shall therefore drop all the smaller higher order quantities, for example, Λ^{00} and $\Lambda^{0\alpha}$, since $\Lambda^{00} \sim \rho O(\varepsilon^4)$, $\Lambda^{0\alpha} \sim \rho O(\varepsilon^5)$. Since

$$\partial_\alpha \partial^\alpha U = 4\pi\rho_0, \quad \frac{\partial V^\beta}{\partial x^\beta} = \frac{\partial U}{\partial t},$$

we can obtain from (1.190) and (1.169)

$$\left.\begin{aligned} t^{00} &= \rho\,[1 + v^2/2 + \Pi - U/2] - \frac{1}{8\pi}\,\partial_\alpha\,[U\partial^\alpha U] \\ &\quad + \rho O(\varepsilon^4), \\ t^{0\alpha} &= \rho v^\alpha \left[1 + \frac{v^2}{2} + \Pi + U\right] + p v^\alpha + 2\rho V^\alpha \\ &\quad + \frac{1}{4\pi}\left\{\frac{\partial U}{\partial t}\,\partial^\alpha U + 2\partial_\beta\,[U\partial^\alpha V^\beta - V^\beta \partial^\alpha U]\right\} \\ &\quad + \rho O(\varepsilon^5). \end{aligned}\right\} \qquad (1.192)$$

To find out what the system's four-momentum in the post-Newtonian approximation, let us integrate expression (1.192) over the whole space. Using the identities

$$\begin{aligned} &\int \frac{\partial U}{\partial t}\,\partial^\alpha U\,dV = 2\pi \int \rho\,[Uv^\alpha + W^\alpha]\,dV, \\ &\int \rho V^\alpha dV = -\int \rho U v^\alpha\,dV, \\ &\int \partial_\alpha\,(U\partial^\alpha U)\,dV = \int dS_\alpha U \partial^\alpha U = 0, \end{aligned}$$

we will finally get

$$\left.\begin{aligned} P^0 &= \int dV \rho \left[1 + \Pi + \frac{v^2}{2} - \frac{1}{2}\,U\right], \\ P^\alpha &= \int dV \left\{\rho v^\alpha \left[1 + \Pi + \frac{v^2}{2} - \frac{1}{2}\,U\right]\right. \\ &\quad \left. + p v^\alpha + \frac{1}{2}\,\rho W^\alpha\right\}. \end{aligned}\right\} \qquad (1.193)$$

Using (1.193) and (1.188) the conserved angular momentum tensor of the system is readily available for the post-Newtonian approximation:

$$\left.\begin{aligned} M^{0\alpha} &= -\int dV \rho x^\alpha \left[1 + \Pi + \frac{v^2}{2} - \frac{1}{2}\,U\right] + P^\alpha t, \\ M^{\alpha\beta} &= \int dV \rho \left\{x^\alpha \left[v^\beta \left(1 + \Pi + \frac{v^2}{2} - \frac{1}{2}\,U + \frac{p}{\rho}\right)\right.\right. \\ &\quad \left. + \frac{1}{2}\,W^\beta\right] - x^\beta \left[v^\alpha \left(1 + \Pi + \frac{v^2}{2} - \frac{1}{2}\,U + \frac{p}{\rho}\right)\right. \\ &\quad \left.\left. + \frac{1}{2}\,W^\alpha\right]\right\}, \end{aligned}\right\} \qquad (1.194)$$

and the coordinates of the system's center-of-mass

$$X^\alpha = \int dV \rho x^\alpha \left[1 + \Pi + \frac{v^2}{2} - \frac{1}{2} U\right] \Big/ P^0. \tag{1.195}$$

Notice that when the accepted system of units is used, the expression for the four-momentum component P^0 (1.187) of an isolated system coincides with that for the inertial mass of this system. Hence we have for the inertial mass in the post-Newtonian approximation

$$m = \int dV \rho \left[1 + \Pi + \frac{v^2}{2} - \frac{1}{2} U\right]. \tag{1.196}$$

The post-Newtonian expression for the inertial mass of a static, spherically symmetric body looks like

$$m = 4\pi \int_0^\alpha r^2 dr \rho \left[1 + \Pi - \frac{1}{2} U\right]. \tag{1.197}$$

The expressions obtained for the inertial (1.197) and gravitational (1.182) masses enable the numerical value of the parameter β to be calculated for the field theory of gravitation. In fact the condition that these masses are equal, which follows from (1.182) and (1.197), unambiguously means that the value of the parameter $\beta = 1$.

Contrary to conservation law (1.185), the covariant equation of conservation of the energy-momentum tensor density of matter in Riemannian space-time,

$$\nabla_n T^{ni} = \partial_n T^{ni} + \Gamma^i_{nm} T^{nm} = 0, \tag{1.198}$$

does not explicitly conserve any quantity but rather mirrors the fact that the energy-momentum tensor density of matter is not conserved,

$$\partial_n T^{ni} \neq 0.$$

However, as has been shown in Sec. 1.4 of the present work, conservation law (1.185) and conservation equation (1.198) are simply the different forms of the same conservation law in the field theory of gravitation. This general result, obtained in Sec. 1.4, may be also confirmed at any stage of the approximation. By the field theory of gravitation therefore integrals of motion (1.193) and (1.194) may be also obtained from conservation equation (1.198) in the post-Newtonian approximation.

We will show, for example, that in the field theory of gravitation the post-Newtonian integrals of motion, that are obtained from covariant equation (1.198) in Riemannian space-time, coincide with the integrals of motion (1.193) that are obtained from conservation law (1.185) in pseudo-Euclidean space-time. It is to be stressed that in order to correlate the integrals of motion the computations must

be carried out in the same coordinate frame for both cases, since different coordinate frames result in different expressions for the integrals of motion.

For this reason the computations will be performed in the "non-canonical" coordinate frame of Riemannian space-time, for which the metric tensor has the form (1.176). This coordinate frame corresponds to the coordinate frame in pseudo-Euclidean space-time for which we obtained the integrals of motion (1.193) and (1.194). Also notice that the transformation to the canonical coordinate frame, whose metric tensor g_{ni} has the form (1.179), can be shown to leave the post-Newtonian expressions for the integrals of motion in the field theory of gravitation unchanged. Yet in the general case different coordinate frames will have different expressions for the integrals of motion.

Using the post-Newtonian expansion for the metric tensor (1.176) and definition (1.153), we can define the components of the energy-momentum tensor density for matter with a post-Newtonian degree of accuracy:

$$\left.\begin{aligned} T^{00} &= \rho\left[1+\Pi+\frac{1}{2}v^2+U\right]+\rho O(\varepsilon^4),\\ T^{0\alpha} &= \rho v^\alpha\left[1+\Pi+\frac{1}{2}v^2+U\right]+pv^\alpha+\rho O(\varepsilon^5),\\ T^{\alpha\beta} &= \rho v^\alpha v^\beta\left[1+\Pi+\frac{1}{2}v^2+U\right]\\ &+pv^\alpha v^\beta-p\gamma^{\alpha\beta}+\rho O(\varepsilon^6). \end{aligned}\right\} \tag{1.199}$$

Let us write equations (1.198) in components form:

$$\left.\begin{aligned} &\partial_0 T^{00}+\partial_\alpha T^{0\alpha}+\Gamma^0_{00}T^{00}+2\Gamma^0_{0\alpha}T^{0\alpha}+\Gamma^0_{\alpha\beta}T^{\alpha\beta}=0,\\ &\partial_0 T^{0\alpha}+\partial_\beta T^{\alpha\beta}+\Gamma^\alpha_{00}T^{00}+2\Gamma^\alpha_{0\beta}T^{0\beta}+\Gamma^\alpha_{\beta\varepsilon}T^{\beta\varepsilon}=0. \end{aligned}\right\} \tag{1.200}$$

Since the components of the energy-momentum tensor density of matter are known with the following accuracies:

$$T^{00}\sim\rho O(\varepsilon^4),\quad T^{0\alpha}\sim\rho O(\varepsilon^5),\quad T^{\alpha\beta}\sim\rho O(\varepsilon^6),$$

the Riemannian space-time relations have to be determined to the following accuracies:

$$\begin{aligned} &\Gamma^0_{00}\sim O(\varepsilon^5),\quad \Gamma^0_{0\alpha}\sim O(\varepsilon^4),\quad \Gamma^0_{\alpha\beta}\sim O(\varepsilon^3),\\ &\Gamma^\alpha_{00}\sim O(\varepsilon^6),\quad \Gamma^\alpha_{0\beta}\sim O(\varepsilon^5),\quad \Gamma^\alpha_{\beta\varepsilon}\sim O(\varepsilon^4). \end{aligned}$$

By using the post-Newtonian expansion of the metric (1.176) we can determine the relations in Riemannian space-time with the required accuracies:

$$\left.\begin{aligned}
&\Gamma^0_{00} = -\partial U/\partial t + O(\varepsilon^5), \quad \Gamma^0_{0\alpha} = -\partial_\alpha U + O(\varepsilon^4),\\
&\Gamma^\alpha_{\beta\varepsilon} = \delta^\alpha_\varepsilon \partial_\beta U + \delta^\alpha_\beta \partial_\varepsilon U - \gamma_{\beta\varepsilon}\partial^\alpha U + O(\varepsilon^4),\\
&\Gamma^\alpha_{0\beta} = \delta^\alpha_\beta\, \partial U/\partial t + 2\gamma^{\alpha\varepsilon}(\partial V_\varepsilon/\partial x^\beta - \partial V_\beta/\partial x^\varepsilon)\\
&+ O(\varepsilon^5), \quad \Gamma^0_{\alpha\beta} = O(\varepsilon^3),\\
&\Gamma^\alpha_{00} = 4\partial V^\alpha/\partial t + \gamma^{\alpha\beta}(1-2U)\,\partial U/\partial x^\beta\\
&-\frac{\gamma^{\alpha\beta}}{2}\frac{\partial}{\partial x^\beta}[2\beta U^2 - 4\Phi_1 + 4(\beta-2)\Phi_2 - 2\Phi_3\\
&-6\Phi_4 - w] + O(\varepsilon^6),
\end{aligned}\right\} \quad (1.201)$$

where

$$w = \frac{\partial^2}{\partial t^2} Q, \quad Q = \int dV \rho R.$$

Substituting expressions (1.199) and (1.201) into the first equation of (1.200), we obtain

$$\frac{\partial}{\partial t}\left[\rho\left(1 + \frac{v^2}{2} + \Pi + U\right)\right] + \partial_\alpha\left[\rho v^\alpha\left(1 + \Pi + U + \frac{v^2}{2}\right)\right.$$
$$\left. + p v^\alpha\right] - \rho\frac{\partial U}{\partial t} - 2\rho v^\alpha \partial_\alpha U = \rho O(\varepsilon^5). \quad (1.202)$$

The second equation of (1.200), allowing for expressions (1.201), becomes

$$\frac{\partial}{\partial t}\left[\rho v^\alpha\left(1 + \Pi + \frac{v^2}{2} + U\right) + p v^\alpha\right]$$
$$+ \partial_\beta\left[\rho v^\alpha v^\beta\left(1 + \Pi + U + \frac{v^2}{2}\right) + p v^\alpha v^\beta - p\gamma^{\alpha\beta}\right]$$
$$+ \rho\left(1 + \Pi + U + \frac{v^2}{2}\right)\partial^\alpha U + 4\rho\frac{\partial V^\alpha}{\partial t}$$
$$-(2+2\beta)\rho U\,\partial^\alpha U + \rho\partial^\alpha[2\Phi_1 - 2(\beta-2)\Phi_2 + \Phi_3 + 3\Phi_4 + w/2]$$
$$+ 2\rho v^\alpha\frac{\partial U}{\partial t} + 4\rho v^\beta(\partial_\beta V^\alpha - \partial^\alpha V_\beta)$$
$$+ 2\rho v^\alpha v^\beta \partial_\beta U + \rho v^2 \partial^\alpha U + p\partial^\alpha U = \rho O(\varepsilon^6). \quad (1.203)$$

To simplify these expressions, we can use the continuity equation of a perfect liquid,

$$\frac{\partial\rho}{\partial t} + \frac{\partial}{\partial x^\alpha}(\rho v^\alpha) = 0,$$

$$\rho = \rho_0\left[1 + 3U + \frac{1}{2}v^2 + O(\varepsilon^4)\right],$$

the Newtonian equations of motion of an elastic body,

$$\rho\, dv^{\alpha}/dt = \gamma^{\alpha\beta}(-\rho\, \partial U/\partial x^{\beta} + \partial p/\partial x^{\beta}),$$
$$\rho\, d\Pi/dt = -p\partial_{\alpha} v^{\alpha},$$

and the relations

$$\frac{\partial V_{\alpha}}{\partial x^{\beta}} - \frac{\partial V_{\beta}}{\partial x^{\alpha}} = \frac{\partial W_{\alpha}}{\partial x^{\beta}} - \frac{\partial W_{\beta}}{\partial x^{\alpha}},$$
$$\partial_{\alpha}\partial^{\alpha} U = 4\pi\rho_0,$$
$$\rho\frac{\partial U}{\partial t} = \frac{1}{2}\frac{\partial}{\partial t}(\rho U) + \frac{1}{8\pi}\frac{\partial}{\partial x^{\alpha}}\left[\frac{\partial U}{\partial t}\partial^{\alpha} U - \gamma^{\alpha\beta} U \frac{\partial^2 U}{\partial t\, \partial x^{\beta}}\right].$$

As a result of these transformations, we can obtain from expressions (1.202)

$$\begin{aligned}&\frac{\partial}{\partial t}\left[\rho\left(1+\Pi-\frac{1}{2}U+\frac{v^2}{2}\right)\right]\\ &+\partial_{\alpha}\left[\rho v^{\alpha}\left(1+\Pi+U+\frac{v^2}{2}\right)+pv^{\alpha}-\frac{1}{8\pi}\frac{\partial U}{\partial t}\partial^{\alpha} U\right.\\ &\left.+\frac{1}{8\pi}\gamma^{\alpha\beta} U\frac{\partial^2 U}{\partial t\,\partial x^{\beta}}\right]=\rho O(\varepsilon^4).\end{aligned} \tag{1.204}$$

Expression (1.203) reduces to the form

$$\begin{aligned}&\frac{\partial}{\partial t}[\rho v^{\alpha}(1+\Pi+v^2/2+U)+pv^{\alpha}]\\ &+\partial_{\beta}[\rho v^{\alpha}v^{\beta}(1+\Pi+v^2/2+U)+pv^{\alpha}v^{\beta}-p\gamma^{\alpha\beta}]\\ &+4\rho\, dV^{\alpha}/dt+2\rho\frac{d}{dt}(Uv^{\alpha})+\rho_0\gamma^{\alpha\beta}\frac{\partial U}{\partial x^{\beta}}+(4-2\beta)\rho U\partial^{\alpha} U\\ &+\rho(\Pi+2v^2)\partial^{\alpha} U+3p\partial^{\alpha} U+3\rho\gamma^{\alpha\beta}\partial\Phi_4/\partial x^{\beta}-4\rho v^{\beta}\partial^{\alpha} V_{\beta}\\ &+\rho\partial^{\alpha}[2\Phi_1+\Phi_3-2(\beta-2)\Phi_2+w/2]=\rho O(\varepsilon^6).\end{aligned} \tag{1.205}$$

Let us integrate these expressions over the entire space. To begin with, note that

$$\left.\begin{aligned}&\int dV\rho_0\left[\Pi\frac{\partial U}{\partial x^{\beta}}+\frac{\partial\Phi_3}{\partial x^{\beta}}\right]=0,\\ &\int dV\rho_0\left[U\frac{\partial U}{\partial x^{\beta}}+\frac{\partial\Phi_2}{\partial x^{\beta}}\right]=0,\\ &\int dV\left[p\frac{\partial U}{\partial x^{\beta}}+\rho_0\frac{\partial\Phi_4}{\partial x^{\beta}}\right]=0,\\ &\int dV\rho_0\left[v^2\frac{\partial U}{\partial x^{\beta}}+\frac{\partial\Phi_1}{\partial x^{\beta}}\right]=0.\end{aligned}\right\} \tag{1.206}$$

To verify these, consider, for example, the first relation. Using equations (1.148), we have

$$I=\int dV\,dV'\rho_0\rho_0'\left[\frac{\Pi\,(x_\beta-x'_\beta)}{|x-x'|^3}+\frac{\Pi'\,(x_\beta-x'_\beta)}{|x-x'|^3}\right].$$

Since the integrand is antisymmetric with respect to the substitution $\rho_0 \leftrightarrow \rho_0'$, $x_\beta \leftrightarrow x'_\beta$, this integral will vanish. The remaining relations in (1.206) can be proved in an analogous fashion.

In addition, we can use the apparent identities

$$\int \rho_0 v^\varepsilon \frac{\partial V_\varepsilon}{\partial x^\beta}\,dV=\int \rho_0 v^\varepsilon \frac{\partial W_\varepsilon}{\partial x^\beta}\,dV=0,$$

$$\int \rho_0 \frac{\partial U}{\partial x^\beta}\,dV=0,\quad \int \rho v^\beta \frac{\partial^3 Q}{\partial t\,\partial x^\alpha\,\partial x^\beta}\,dV=0,$$

$$\frac{d}{dt}\int \rho f\,dV=\int dV\rho\left[\frac{d}{dt}f+fO(\varepsilon^2)\right].$$

Again, we can use the fact that the volume integrals of spatial divergence vanish when they are transformed into surface integrals.

As a result, we can obtain an energy integral $dP^0/dt=0$ from expression (1.204) for the post-Newtonian approximation of the field theory of gravitation, where

$$P^0=\int dV\rho\left[1+\Pi+\frac{1}{2}v^2-\frac{1}{2}U\right]=\text{const.}\tag{1.207}$$

Allowing for

$$\rho\partial^\alpha\frac{\partial^2}{\partial t^2}Q=\frac{\partial}{\partial t}\left(\rho\partial^\alpha\frac{\partial}{\partial t}Q\right)$$
$$+\partial_\beta\left[\rho v^\beta\partial^\alpha\frac{\partial}{\partial t}Q\right]-\rho v^\beta\partial^\alpha\partial_\beta\frac{\partial}{\partial t}Q,$$

we obtain from (1.205)

$$\frac{dP^\alpha}{dt}=0.$$

Whence it follows that

$$P^\alpha=\int dV\left[\rho v^\alpha\left(1+\Pi+\frac{1}{2}v^2+3U\right)\right.$$
$$\left.+pv^\alpha+4\rho V^\alpha+\frac{1}{2}\partial^\alpha\frac{\partial}{\partial t}Q\right]=\text{const.}\tag{1.208}$$

Inasmuch as the following relations are valid:

$$\partial^\alpha\frac{\partial}{\partial t}Q=W^\alpha-V^\alpha,$$

$$\int dV\rho V^\alpha=-\int dV\rho U v^\alpha,$$

and expression (1.208) may be reduced to

$$P^{\alpha}=\int dV\rho\left[v^{\alpha}\left(1+\Pi+\frac{v^2}{2}-\frac{1}{2}U+\frac{p}{\rho}\right)+\frac{1}{2}W^{\alpha}\right]. \tag{1.209}$$

Thus, in the field theory of gravitation, the post-Newtonian integrals of motion obtained from conservation law (1.185) in pseudo-Euclidean space-time coincide with those from covariant equation (1.198) in Riemannian space-time. This is a direct corollary of the fact that in the field theory of gravitation conservation law (1.185) and equation (1.198) are the different forms of the same conservation law. In the field theory a gravitational field is a physical one possessing an energy-momentum density and contributing to the complete energy-momentum tensor of the system. It is the presence of the conventional conservation laws in the field theory of gravitation which makes it possible to carry out different energy-based computations which include finding the post-Newtonian expressions for the integrals of motion.

The GTR field is not a Faraday-Maxwell field with the result that in Einstein's theory there is no way to calculate a gravitational field's energy. Yet GTR post-Newtonian integrals of motion of an isolated system can generally be obtained from the covariant equation (1.198) arriving at expressions (1.207) and (1.209).

The reason for this is easy to comprehend, since we can obtain from the covariant conservation law (1.198) in the GTR either conservation law (1.36) which holds trivially by virtue of Einstein's equations or the relation containing the energy-momentum pseudotensor:

$$\partial_i\left[-g\left(T^{ik}+\tau^{ik}\right)\right]=0. \tag{1.210}$$

In the field theory of gravitation, conservation equation (1.198) is equivalent to the conservation law for the energy-momentum tensor of matter and a gravitational field when taken in conjunction:

$$D_i\left[t_{\mathrm{g}}^{ik}+t_{\mathrm{M}}^{ik}\right]=0. \tag{1.211}$$

The GTR enables integral quantities to be obtained from relation (1.210), which are time independent in the lowest approximation:

$$J^{k}=\int dV\,(-g)\left[T^{0k}+\tau^{0k}\right]. \tag{1.212}$$

These quantities cannot, however, be interpreted as integrals of motion. Since they reflect the choice of coordinates rather than the physical characteristics of a system which is composed of matter and a gravitational field, they can always be made equal to any prescribed values, both positive and negative, by an appropriate choice of coordinates leaving the metric tensor of Riemannian space-time asymptotically Galilean. As a consequence, quantities (1.212)

are not the GTR integrals of motion. Yet in the first approximation, in Cartesian coordinates, expressions (1.210) and (1.211) coincide. This is why the expressions for the quantities (1.212) coincide in this coordinate frame with expression (1.187) for the integrals of motion of the field theory of gravitation.

It is because a theory of gravitation can be constructed in pseudo-Euclidean space-time to possess the tensorial conservation law for the energy-momentum of matter and gravitational field when taken in conjunction that in the first approximations quantities (1.212) yield, in a narrow class of coordinate frames, results which coincide with expressions (1.193) for the integrals of motion of the field theory of gravitation. However, as a result of their different transformation laws, the GTR expressions (1.212) would differ substantially in another frame while quantities (1.187) in the field theory of gravitation would conserve their physical meaning as integrals of motion. Since the GTR quantities (1.212) do not have any physical meaning as integrals of motion, physically absurd results can be easily obtained when expression (1.212) is in another coordinate frame. For this reason the generally accepted interpretation of them in Einstein's theory, as the energy-momentum of an isolated system, is wrong. We notice as we conclude this section that the energy of a static field, computed from the Newtonian approximation for the field theory of gravitation, using the canonical energy-momentum tensor (1.101), is positive, viz.

$$\int \widetilde{t}_{\mathrm{g}}^{00}\, dV = -\frac{1}{8\pi}\int dV \partial_\alpha U \partial^\alpha U > 0.$$

However, when computed using the symmetric energy-momentum tensor (1.105), the energy is negative:

$$\int t_{\mathrm{g}}^{00}\, dV = \frac{3}{8\pi}\int dV \partial_\alpha U \partial^\alpha U < 0.$$

As is generally known, the opposite is true in electrodynamics. Electrostatic field energy, computed using a canonical energy-momentum tensor, is negative, while the same energy computed via a symmetric tensor is positive. By analogy we can conclude that a static gravitational field is attractive, since in electrodynamics charges of the same sign produce a repulsive field.

When the total energy of the matter and a static gravitational field is calculated for the Newtonian approximation, the same result is obtained whether the canonical or symmetric energy-momentum tensor is used, viz.

$$P^0 = \int dV\, [t_{\mathrm{g}}^{00} + t_{\mathrm{M}}^{00}] = \int dV\, [\widetilde{t}_{\mathrm{g}}^{00} + \widetilde{t}_{\mathrm{M}}^{00}]$$
$$= \int dV \rho \left[1 + \Pi + \frac{v^2}{2} - \frac{U}{2}\right].$$

It follows from this expression that the energy of two rest particles increases with the separation distance. This too exhibits the action of mutual attraction forces.

1.13 GRAVITATIONAL EXPERIMENTS IN THE SOLAR SYSTEM

Consider the limitations imposed on the values of the post-Newtonian parameters by experiments performed within the solar system.

We will analyse these experiments thus: at the outset we shall consider the standard effects, the deflection of light and radio waves in the Sun's field, the advance of Mercury's perihelion, and the measurement of the time delay of a radio signal in the Sun's gravitational field. Then we will consider the Nordtvedt effect and those that occur because the parameters α_1, α_2, α_3, and ξ_w do not vanish. Red shift in the Sun's gravitational field will not be considered here since this effect can be completely described in the Newtonian approximation [87].

1.13.1 The Deflection of Light and Radio Waves in the Sun's Gravitational Field

According to [88], light and radio waves which are considered as massless particles having an impact parameter b are deflected in the Sun's gravitational field through an angle

$$\delta\varphi = \frac{2(1+\gamma)M}{b}.$$

Analysis of the experimental results obtained by observing the way the light from remote stars and the radio waves emitted by quasars are bent in the Sun's gravitational field gives ground for the assumption [72] that the post-Newtonian parameter γ is

$$\gamma = 1.0 \pm 0.2.$$

1.13.2 The Time Delay of Radio Signals in the Sun's Field

Another independent way of determining the post-Newtonian parameter γ is to measure the time delay of a radio signal in the Sun's field [89].

This effect is due to the fact that the propagation time of radio signals sent from the Earth to a reflector located in some other part of the solar system and back again measured by a clock located on the Earth is different from the time the same process would take if the gravitational field were absent. The surface of planets and the

radio equipment on artificial satellites have been used as reflectors for these experiments.

These experimen ts [90] resulted in

$$\gamma = 1.000 \pm 0.002.$$

The field theory of gravitation, as does Einstein's GTR, yields for the parameter γ the result $\gamma = 1$ and this agrees with the results of experiments.

1.13.3 Precession of an Orbiting Gyroscope

If the parameters $\alpha_1 = \alpha_2 = \alpha_3 = 0$, then measuring the precession of a gyroscope orbiting around the Earth will be a third, independent method of determining the parameter γ. From [91] the velocity of the angular precession of a gyroscope placed in a circular orbit in Earth's field will be equal to

$$\mathbf{\Omega} = \frac{2\gamma+1}{2} m \frac{\mathbf{r} \times \mathbf{v}}{r^3} + \frac{\gamma+1}{2r^3}\left(-\mathbf{J} + \frac{3\mathbf{r}(\mathbf{r}\cdot\mathbf{j})}{r^2}\right),$$

where m is Earth's mass, $\mathbf{v}$ the linear velocity of the gyroscope relative to the Earth's center, $\mathbf{J}$ Earth's angular momentum, and $\mathbf{r}$ the radius-vector of a point where the gyroscope is located.

Modern technological developments [92-94] raise hopes that this experiment will be performed in the near future.

1.13.4 The Advance of Mercury's Perihelion

The magnitude of the advance of Mercury's perihelion is influenced by a number of factors other than the occurrence of post-Newtonian corrections to the equations of motion. These include the attraction from other planets in the solar system, and the availability of the Sun's quadrupole moment. The only uncertain factor among them is the magnitude of the Sun's quadrupole moment as the influence of all the other factors can be computed with an adequate degree of accuracy.

The overall advance of Mercury's perihelion due to the Sun's quadrupole moment J_2 and to the post-Newtonian corrections of the equation of motion from theoretical considerations [95] is given by

$$\delta\varphi = 42.98\left[\frac{2+2\gamma-\beta}{3}\right] + 1.3\cdot 10^5 J_2$$

in angular seconds per century. The results of observations [72] imply that $\delta\varphi = 41.4 \pm 0.9$ angular seconds per century.

Measuring the visible shape of the Sun, Dicke and Goldberg [96] have obtained a magnitude for $J_2 = (2.5 \pm 0.2)\cdot 10^{-5}$ and latter measurements by Hill *et al.* [95] have demonstrated that $J_2 < < 0.5\cdot 10^{-5}$. By comparing the perihelion advances observed for

Mercury and Mars [97], we obtain another estimate of the magnitude $J_2 \leqslant 3 \cdot 10^{-5}$.

Thus, the lack of direct measurements of the Sun's quadrupole moment leaves a great deal of uncertainty in the value of β determined from the perihelion advance of Mercury:

$$\beta = 1.0 \begin{cases} +0.4 \\ -0.2. \end{cases}$$

Note that the field theory of gravitation requires that the parameter $\beta = 1$ and this is within the error limits.

1.13.5 The Nordtvedt Effect and the Laser Location of the Moon

It has been shown in a number of works (see [98-100]) that the presence of post-Newtonian corrections to the equations of motion of an extended body will be accompanied by a number of anomalies in the motion of its center-of-mass. One of these effects is the polarization of the lunar orbit towards the Sun. The eccentricity of the orbit, which is pulled towards the Sun due to this polarization, has an amplitude

$$\delta r = C_0 \eta \cos \theta_0,$$

where C_0 is a constant of the order of 10 m, θ_0 the difference between the longitude of the Moon and that of the Sun, and

$$\eta = 4\beta - \gamma - 3 - \alpha_1 + \frac{2}{3}\alpha_2 - \frac{2}{3}\xi_1 - \frac{1}{3}\xi_2 - \frac{10}{3}\xi_w.$$

Measurements of the Moon's location using lasers [101, 102] have shown that the magnitude of $C_0\eta$ is

$$C_0\eta = 0.0 \pm 4.0 \text{ cm},$$

whence it is possible to obtain an estimate of η as

$$\eta = 0.00 \pm 0.03.$$

Thus, these experiments can be described by the field theory of gravitation within the limits of experimental error.

1.13.6 The Effects Associated with the Availability of a Preferred Frame of Reference

The theories of gravitation, for which only one of the parameters, α_1, α_2 or α_3, does not vanish, possess a preferred frame of reference. The predictions these theories make regarding the standard effects can agree with the results of observations only if the solar system is the preferred frame of reference. Yet it is more

reasonable to assume the solar system moves relative to other stellar systems and is not preferred when compared to the latter. It cannot therefore be the preferred universal rest system for the theories under consideration.

Since a preferred rest system must be singled out from the other systems, it would be more judicious to associate the system with the Galactic center-of-mass or even with the Universe center-of-mass. The solar system would then move at the velocity of $\sim 10^{-3}c$ with respect to the preferred rest frame, which is of the same order of magnitude as the orbital velocity of the solar system relative to the Galactic center. In this case it would be possible to observe a number of effects which involve a motion relative to the preferred rest frame [72], and so allow the parameters α_1, α_2, and α_3 to be evaluated.

In theories of gravitation which have a preferred rest frame the gravitational constant G measured in the gravimetric experiments would depend on the Earth's motion with respect to this frame.

Theoretically, [72] would give for the relative quantity $\Delta G/G$

$$\Delta G/G = \left(\frac{\alpha_2}{2} + \alpha_3 - \alpha_1\right) \mathbf{w}\cdot\mathbf{v} + (1/4)\,\alpha_2\,[(\mathbf{v}\cdot\mathbf{e}_r)^2 + 2\,(\mathbf{w}\cdot\mathbf{e}_r)\,(\mathbf{v}\cdot\mathbf{e}_r) + (\mathbf{w}\cdot\mathbf{e}_r)^2],$$

where $\mathbf{v}$ is the orbital velocity of the Earth about the Sun, $\mathbf{w}$ the velocity of the Sun with respect to the preferred rest frame, and $\mathbf{e}_r$ a unit vector directed from the gravimeter to the Earth center.

As a result of the Earth rotation about its axis, the vector $\mathbf{e}_r$ would change its orientation relative to the vectors $\mathbf{v}$ and $\mathbf{w}$ leading to a periodic change in the scalar products $\mathbf{v}\cdot\mathbf{e}_r$ and $\mathbf{w}\cdot\mathbf{e}_r$ with a period of about 12 hr. This would result in corresponding periodic changes in the magnitude of free fall acceleration. Thus for a point of observation located at the latitude θ we would have

$$\frac{\Delta g}{g} \approx 3\alpha_2\cdot 10^{-8}\cos^2\theta$$

Analyzing the results of gravimetric experiments, Will [103] has found that the relative changes in the magnitude of g do not exceed 10^{-9}, viz. $\left|\frac{\Delta g}{g}\right| < 10^{-9}$. Whence an estimate of the value of α_2 can be obtained: $|\alpha_2| < 3\cdot 10^{-2}$.

The Earth's motion about the Sun is also accompanied by the periodic changes in the magnitude of $\mathbf{w}\cdot\mathbf{v}$ which would have a period of the order of one year. This variation would bring about a contraction and expansion of the Earth, which would, in turn, lead to periodic changes in the angular velocity of the Earth's rotation due to a change in the moment of inertia:

$$\frac{\Delta\omega}{\omega} \approx 3\cdot 10^{-9}\left(\alpha_3 + \frac{2}{3}\alpha_2 - \alpha_1\right).$$

It follows from the results of observations that

$$\left|\alpha_3+\frac{2}{3}\alpha_2-\alpha_1\right|<0.2.$$

The solar system motion relative to the center of the Universe may give rise to an anomalous perihelion advance, $\delta\varphi_0$, of planets. The additional contribution to the advance of Mercury's perihelion [72] (in angular seconds per century) would amount to

$$\delta\varphi_0 = 35\alpha_1 + 8\alpha_2 - 4\cdot 10^4\alpha_3.$$

Comparing with observational data and combining all the estimates of the parameter α, we get

$$|\alpha_1| < 0.2, \quad |\alpha_2| < 3\cdot 10^{-2}, \quad |\alpha_3| < 2\cdot 10^{-5}.$$

In the field theory of gravitation, as in Einstein's GTR, $\alpha_1 = \alpha_2 = \alpha_3 = 0$, and therefore all these effects are absent.

1.13.7 The Effects of Anisotropy Relative to the Galactic Center

The effects of anisotropy due to the Galaxy's gravitational field [81] are only conceivable for theories of gravitation where the parameter ξ_w does not vanish.

If the mass M of the Galaxy is assumed to be concentrated at the Galaxy center, which is at a distance R from tne solar system, the galactic gravitational field would cause periodic changes in gravimeter readings having a period of 12 hr:

$$\frac{\Delta G}{G}=\xi_w\left(1-\frac{3K}{mr^2}\right)\frac{M}{R}(\mathbf{e}_r\cdot\mathbf{e}_R),$$

where K is the moment of inertia, m Earth's mass and r its radius, $\mathbf{e}_r$ a unit vector directed from the gravimeter to the Earth's center, and $\mathbf{e}_R = \mathbf{R}/R$.

Another effect would be an anomalous perihelion advance of the planets, conditioned by the anisotropy due to the Galaxy:

$$\delta\varphi_0=\frac{\pi\xi_w}{2}\frac{M}{R}\cos^2\beta\cos^2(\omega-\lambda),$$

where λ and β are the angular coordinates of the Galactic center, ω the perihelion angle in terms of geocentric coordinates.

A comparison of observations yields an upper limit, viz.

$$|\xi_w| < 10^{-2}.$$

For the field theory of gravitation, as for Einstein's GTR, $\xi_w = 0$ and all the effects of anisotropy that would be produced by the Galaxy's gravitational field are absent.

Closing this survey of gravitational experiments, we arrive at the conclusion that the field theory of gravitation enables all the experimental facts to be covered. Notice that in the post-Newtonian limit the quadratic terms of the coupling equation (1.95) are indistinguishable, since no experiment in the gravitational field of the solar system will allow the minimum-coupling parameters to be determined separately at the post-Newtonian level.

As we will show in Sec. 1.17, measuring the deceleration parameter of the expanding Universe in the vicinity of the present time would enable the magnitude of another linear combination of these minimum-coupling parameters to be determined. The minimum-coupling parameters can only be defined by increasing the measurement accuracy of experiments in the gravitational field of the solar system beyond a post-Newtonian level.

Summarizing, we notice that the equivalence principle is valid in the field theory of gravitation for point particles only. For extended bodies moving in a weak gravitational field it is only approximate, and accuracy of the weak gravitational field can be assumed to be homogeneous in the region occupied by the body. In this case the gravitational field can be "eliminated" by passing to a coordinate system where $g_{ni} = \gamma_{ni}$ within the region occupied by matter. It appears from experiments by Braginskii [104] with the bodies of laboratory dimensions in a sufficiently homogeneous gravitational field, that the equivalence principle is valid for strong, electromagnetic, and weak interactions to the accuracy achieved in these experiments. Yet for extended bodies, this principle is rigorously invalid if taking due regard for the gravitational field both in Einstein's GTR and in the field theory of gravitation.

1.14 A STATIC, SPHERICALLY SYMMETRIC GRAVITATIONAL FIELD

In the case of a static source of radius a with a spherically symmetric distribution of matter, the gravitational field equations (1.91) and the expression for the tensor current (1.85) are appreciably simpler.

Proceeding from symmetry of the problem, we can define which components of the tensors I_{lm} and h^{lm} will not vanish in this case. Assume the origin of a spherical coordinate frame is at the center of the source. Rotating this coordinate frame through an arbitrary angle, the physical situation, by virtue of the spherical symmetry of the matter distribution, will not change. The tensor components I_{lm} and h^{lm} for a rotational transformation must, therefore, be the same functions of the transformed arguments, as the initial functions are of their initial arguments, i.e. these tensors must be forminvariant for rotational transformations of the coordinate frame. Whence

it follows that in the spherical coordinate frame only (00), $(0r)$, (rr), $(\theta\theta)$, $(\varphi\varphi)$ could be non-vanishing components of the tensors I_{lm} and h^{lm}, since only it is in this case that the tensors I_{lm} and h^{lm} are forminvariant for rotational transformations.

Expressions (1.85) and (1.74) result in

$$I_{0r} = 0.$$

Hence for the case of a static, spherically symmetric distribution of matter the tensor I_{lm} has the following components:

$$I_{lm} = \{I_{00}, I_{rr}, I_{\theta\theta}, I_{\varphi\varphi} = I_{\theta\theta} \sin^2 \theta\}.$$

In this case the gravitational field equations (1.19) are written as a system of ordinary second-order differential equations:

$$\left.\begin{aligned} & f''_{00} + \frac{2}{r} f'_{00} = 16\pi I_{00}(r), \\ & f''_{rr} + \frac{2}{r} f'_{rr} - \frac{4}{r^2} f_{rr} + \frac{4}{r^2}\left(\frac{f_{\theta\theta}}{r^2}\right) = 16\pi I_{rr}(r), \\ & \left(\frac{f_{\theta\theta}}{r^2}\right)'' + \frac{2}{r}\left(\frac{f_{\theta\theta}}{r^2}\right)' - \frac{2}{r^2}\left(\frac{f_{\theta\theta}}{r^2}\right) + \frac{2}{r^2} f_{rr} \\ & = 16\pi \frac{I_{\theta\theta}(r)}{r^2}. \end{aligned}\right\} \tag{1.213}$$

Here and below a prime denotes derivatives with respect to r.

As the boundary conditions for these equations we shall require the function f_{00}, f_{rr} and $f_{\theta\theta}/r^2$ to be bounded at $r = 0$ and to vanish at $r \to \infty$. Then the solution of the gravitational field equations (1.213) will be unique.

However, the components of the tensor current in expression (1.213) are not independent because of the conditions $D^l I_{lm} = 0$. In our case these conditions take the form

$$I'_{rr} + \frac{2}{r}\left(I_{rr} - \frac{1}{r^2} I_{\theta\theta}\right) = 0. \tag{1.214}$$

Expressing component $I_{\theta\theta}$ from this equation, we substitute it into relation (1.213). Integrating the equations and taking into account that $I_{rr} = 0$ outside the source, the gravitational field components become

$$\left.\begin{aligned} & f_{00} = -16\pi\left\{\frac{1}{r}\int_0^r r_0^2\, dr_0 I_{00} + \int_r^a r_0\, dr_0 I_{00}\right\}, \\ & f_{rr} = -\frac{16\pi}{3}\left\{\frac{1}{r^3}\int_0^r r_0^4\, dr_0 I_{rr} + \int_r^a \frac{dr_0}{r_0} I_{rr}\right\}, \\ & f_{\theta\theta} = -\frac{16\pi}{3}\left\{-\frac{1}{2r^3}\int_0^r r_0^4\, dr_0 I_{rr} + \int_r^a \frac{dr_0}{r_0} I_{rr}\right\}. \end{aligned}\right\} \tag{1.215}$$

Consider an external ($r > a$) solution. By introducing the quantities

$$M = 4\pi \int_0^a r_0^2 \, dr_0 I_{00}, \quad \mu = \frac{4\pi}{3} \int_0^a r_0^4 \, dr_0 I_{rr}, \tag{1.216}$$

for the external solution we obtain the following expressions:

$$f_{00} = -\frac{4M}{r}, \quad f_{rr} = -\frac{4\mu}{r^3}, \quad f_{\theta\theta} = \frac{2\mu}{r}, \quad f_{\varphi\varphi} = f_{\theta\theta} \sin^2\theta. \tag{1.217}$$

As has been noted in Sec. 1.7, the f_{lm} fields may be subjected to the gauge transformation,

$$f_{lm} \rightarrow f_{lm} + D_l a_m + D_m a_l - \gamma_{lm} D_n a^n, \tag{1.218}$$

with a gauge vector a^n that satisfies the equation $D_l D^l a^n = 0$. Under this transformation the Lagrangian density of the gravitational field can only be changed by a four dimensional divergence, though this is not essential for the theory. A change in the Riemannian space-time metric tensor g_{ni} produced under the transformation (1.218), corresponds to a coordinate transformation of Riemannian space-time and may invariably be eliminated via an appropriate choice of the coordinates.

We can use gauge transformation (1.218) to simplify the external solution (1.217). Due to the symmetry of the problem under consideration, the gauge vector a_n satisfying the condition $D_l D^l a_n = 0$ can be chosen to have the form $a_r = -\mu/r^2$, $a_0 = a_\theta = a_\varphi = 0$. Then, as a result of this gauge transformation, the following external solution is produced:

$$f_{00} = -4M/r, \quad f_{rr} = f_{\theta\theta} = f_{\varphi\varphi} = 0. \tag{1.219}$$

To obtain the metric tensor when the source is static and spherically symmetric, the gravitational field components (1.219) must be substituted into minimum-coupling equations (1.95). As a result, we obtain (for $r > a$)

$$g_{00} = 1 - 2M/r + 2M^2/r^2, \quad g_{\alpha\beta} = \gamma_{\alpha\beta} [1 + 2M/r + 4\lambda M^2/r^2], \tag{1.220}$$

where $\lambda = b_3 + b_4$.

It can be verified quite easily that the component g_{00} of the effective Riemannian space-time metric tensor (1.220) has no physical singularity outside the source:

$$g_{00} \neq 0, \quad |g_{00}| < \infty$$

for $r > a$.

For the spatial components $g_{\alpha\beta}$ of the metric tensor (1.220) to have no physical singularity outside the gravitational field source too, the following condition has to be satisfied:

$$\lambda = b_3 + b_4 \geqslant 0. \qquad (1.221)$$

From relations (1.216) and (1.85) we have

$$M = 8\pi \int_0^a r_0^2\, dr_0 \left[I_{00} - \frac{1}{2} I_n^n \right] = 8\pi \int_0^a r_0^2\, dr_0 \left[h_{00} - \frac{1}{2} h_n^n \right].$$

For the post-Newtonian expansion of the quantity M to be obtained, the calculations must be carried out, as usual, by successive stages. First, obtain an expansion for M from a Newtonian approximation totally neglecting the influence of gravity on the energy-momentum tensor of matter, and next, use the Newtonian approximation to find a post-Newtonian expression. As a result we obtain

$$M = 4\pi \int_0^a r^2\, dr \{\rho [1 + \Pi - U/2 + O(\varepsilon^4)]\}.$$

As might be expected, the post-Newtonian expansion of the total mass of a static, spherically symmetric body coincides with expression (1.197).

1.15 A NEW MECHANISM OF ENERGY RELEASE IN ASTRONOMICAL OBJECTS

Since the effective Riemannian space-time metric tensor for the case of a static, spherically symmetric source differs appreciably in the field theory of gravitation from the GTR Schwarzschild solution, the description of phenomena taking place in strong gravitational fields must be different in these theories. This means that a number of new effects that occur in the field theory of gravitation and which differ in principle from those in the GTR can be studied in strong gravitational fields.

One of them [105] is a new mechanism for the release of energy in astronomical objects. It can be understood from the following simple reasoning. Using expression (1.220) for the effective Riemannian space-time metric tensor we can find an expression for the force acting on a test mass M_0 at rest outside the source ($r \geqslant a$) and emanating from the static, spherically symmetric field source. The radial component of this force is

$$F^r = -\frac{M_0 m (1 - 2m/r)}{r^2 [1 + 2m/r + 4\lambda m^2/r^2]}, \qquad (1.222)$$

where m is the inertial mass of the source.

It follows from this expression that at $m/r < 1/2$ the force acting on the test mass will be attractive, whereas at $m/r > 1/2$ it will be repulsive. Thus, in the field theory of gravitation with minimum coupling, as the magnitude of the potential increases, gravitationally attractive forces become gravitationally repulsive. This is a radically new property of gravitational interaction which differs substantially from the gravitational properties in the GTR. In particular, it follows from this that in the field theory a collapse is impossible.

However great the gravitational forces that compress an astronomical object may be, the contraction must inevitably come to a stop when the object's dimensions are close to the value of its Schwarzschild radius. At this stage an expansion of the matter will take place accompanied by a partial release of the object's mass. In addition, static astronomical objects with $m/a \gtrsim 1/2$ will be in a state of unstable equilibrium from which they will sooner or later pass into a stable static state with $m/a \ll 1/2$ having released a certain amount of their mass naturally accompanied by a release of part of the object's internal energy in the form of radiation.

The above begs the following questions. At which average value of gravitational potential are astronomical objects of some type (giant stars, supermassive stellar aggregates, etc.) in unstable equilibrium? In what way can these objects be found in that state?

To answer these questions in a rigorous way, we should choose a model of some astronomical object, and then construct an internal structure model of the object by the simultaneous solution of the gravitational field equations and the equation of motion of the matter, incorporating an enthropy variational equation and an equation of state for the matter. The model can then be tested for stability with respect to various perturbations (a random perturbation of the object's radius, a small change in its mass due to the accretion of surrounding matter, a burn-out of matter in its interior, etc.) and so the above questions might be answered. However, the problem is not amenable to analytic solution and requires numerical computer calculations, and this presents substantial difficulties in the analysis of the question.

It is simpler to obtain qualitative order of magnitude estimates, since we can make use of a well-known scheme of analytical investigation for stability in astronomical objects [106]. This analysis requires the use of an averaged equations for the perturbations of an astronomical object which is nearly in a static state. To simplify, consider the spherically symmetric case. As a model of the object's matter, we shall consider a perfect liquid which has an energy-momentum tensor in the effective Riemannian space-time in the form

$$T^{ni} = (\varepsilon + p) u^n u^i - p g^{ni}, \tag{1.223}$$

where $u^i = dx^i/ds$ is the velocity four-vector, ε the mass density, and p the isotropic pressure.

In the field theory of gravitation, as well as in any other metric theory, the energy-momentum tensor of matter satisfies the covariant conservation equation:

$$\nabla_n T^{ni} = 0. \tag{1.224}$$

Let us project this equation on the direction of the velocity four-vector u^i and on its orthogonal direction. As a result, we obtain the covariant continuity equation of the perfect liquid,

$$\nabla_n [(\varepsilon + p) u^n] = u^n \nabla_n p, \tag{1.225}$$

and the equations of motion

$$(\varepsilon + p) u^n \nabla_n u^i = (g^{ni} - u^n u^i) \nabla_n p. \tag{1.226}$$

In the case of spherically symmetric motion and matter distribution continuity equation (1.225) becomes

$$\frac{1}{\sqrt{-g}} \frac{\partial}{\partial t} [(\varepsilon + p) \sqrt{-g}\, u^0] + \frac{1}{\sqrt{-g}} \frac{\partial}{\partial r} [(\varepsilon + p) \sqrt{-g}\, u^0 v^r] = u^0 \left[\frac{\partial p}{\partial t} + v^r \frac{\partial p}{\partial r} \right], \tag{1.227}$$

and the equation of motion (1.226) for the radial component of the velocity four-vector is written as

$$(\varepsilon + p) \left[\frac{\partial u^r}{\partial t} + v^r \frac{\partial u^r}{\partial r} \right] = (g^{nr} - u^n u^r) \partial_n p - (\varepsilon + p) [\Gamma^r_{00} u^0 u^0 + 2\Gamma^r_{0r} u^0 u^r + \Gamma^r_{rr} u^r u^r]. \tag{1.228}$$

Consider a spherically symmetric astronomical object that is in a static state and let us investigate this object for its stability to small perturbations of its parameters (mass, radius, matter density, etc.). Such an investigation is within the scope of perturbation theory.

We expand all the quantities in equations (1.227) and (1.228) into power series in small perturbations and limit ourselves to linear perturbational terms only. In the initial unperturbed state every quantity entering equations (1.227), (1.228) will be assumed to be independent of time. Let us take into account that component g_{0r} of the effective Riemannian space-time metric tensor vanishes in the case of a static, spherically symmetric source.

From the zeroth approximation of equation (1.228) we obtain

$$\overset{(0)}{g^{rr}} \partial_r \overset{(0)}{p} - (\overset{(0)}{\varepsilon} + \overset{(0)}{p}) \overset{(0)}{\Gamma^r_{00}} \overset{(0)}{g^{00}} = 0. \tag{1.229}$$

At the same approximation level we can obtain from continuity equation (1.227)

$$\frac{d}{dt}\left[\sqrt{-\overset{(0)}{g}}\,(\overset{(0)}{\varepsilon}+\overset{(0)}{p})\overset{(0)}{u^0}\right]=0. \tag{1.230}$$

Using an approximation where the perturbations are linear, equation of motion (1.228) reduces to

$$\begin{aligned}(\overset{(0)}{\varepsilon}+\overset{(0)}{p})\overset{(0)}{g^{00}}\frac{d^2}{dt^2}\delta r&=\partial_r\overset{(0)}{p}\delta g^{rr}+\overset{(0)}{g^{rr}}\delta\partial_r p\\ &-(\delta\varepsilon+\delta p)\overset{(0)}{\Gamma^r_{00}}\overset{(0)}{g^{00}}-(\overset{(0)}{\varepsilon}+\overset{(0)}{p})\overset{(0)}{g^{00}}\delta\Gamma^r_{00}\\ &-(\overset{(0)}{\varepsilon}+\overset{(0)}{p})\overset{(0)}{\Gamma^r_{00}}\delta g^{00}.\end{aligned} \tag{1.231}$$

Let us average equations (1.229) and (1.231) over the volume of the astronomical object. Using qualitative estimates, a standard practice for the given approach [105], we obtain the following for the averaged components of the metric:

$$\left.\begin{aligned}\overline{g_{00}}&=1-2m/a+2m^2/a^2,\\ \overline{g_{\alpha\beta}}&=\gamma_{\alpha\beta}\,[1+2m/a+4\lambda m^2/a^2].\end{aligned}\right\} \tag{1.232}$$

For the purposes of our investigation, however, these expressions are more conveniently written in terms of the total rest mass M of the particles composing astronomical object rather than in terms of the inertial mass m of the gravitational field source:

$$M=4\pi\int_0^a r^2\,dr\,\sqrt{-g}\,\varepsilon u^0. \tag{1.233}$$

It follows from (1.197) and (1.233) that

$$m=M\left[1-\frac{M}{2a}+O\left(\frac{M^2}{a^2}\right)\right].$$

Therefore, relations (1.232) take the form

$$\begin{aligned}\overline{g_{00}}&=1-2M/a+3M^2/a^2+O(M^3/a^3),\\ \overline{g_{\alpha\beta}}&=\gamma_{\alpha\beta}\,[1+2M/a+(4\lambda-1)M^2/a^2+O(M^3/a^3)].\end{aligned} \tag{1.234}$$

The averaging of equation (1.229) will then lead to the following identity:

$$\frac{\bar p}{a}=(\bar\varepsilon+\bar p)\,\frac{M}{a^2}\,\frac{[1-3M/a+O(M^2/a^2)]}{[1-2M/a+3M^2/a^2]}. \tag{1.235}$$

Likewise we can obtain from equation (1.231), allowing for equations (1.230) and (1.235), the averaged equation for perturbations of a spherically symmetric astronomical object:

$$(\bar{\varepsilon}+\bar{p})\bar{g}^{00}\delta\ddot{a}=\bar{g}^{rr}\,[-\delta\bar{p}/a+(\dot{\bar{p}}/a^2)\,\delta a$$
$$-(\delta\bar{\varepsilon}+\delta\bar{p})\,\bar{\Gamma}_{r,\,00}\,\bar{g}^{00}-(\bar{\varepsilon}+\bar{p})\,\bar{\Gamma}_{r,\,00}\delta\bar{g}^{00}$$
$$-(\bar{\varepsilon}+\bar{p})\,\bar{g}^{00}\delta\Gamma_{r,\,00}]. \qquad (1.236)$$

Consider a spherically symmetric astronomical object of a radius a, that is in a static state. Let us test this state for its stability to small perturbations of its radius. Assume that the whole process runs adiabatically and that the number of particles and hence the total rest mass of all the particles of this object, too, is conserved.

Since the averaging of expression (1.233) yields

$$M=\frac{4\pi}{3}a^3\bar{\varepsilon}\,[1+2M/a+(4\lambda-1)\,M^2/a^2], \qquad (1.237)$$

the requirement of the total rest mass of the particles to be invariant leads to the following relationship between the value of the average mass density $\delta\bar{\varepsilon}$ and the radial perturbation δa of the object:

$$\frac{\delta\bar{\varepsilon}}{\bar{\varepsilon}}=-\frac{3\delta a}{a}\,\frac{[1+M/a]}{[1+2M/a+(4\lambda-1)\,M^2/a^2]} \qquad (1.238)$$

Introducing the mean adiabatic exponent

$$\bar{\Gamma}_1=\frac{\bar{\varepsilon}}{\bar{p}}\left(\frac{\partial\bar{p}}{\partial\bar{\varepsilon}}\right)_{\text{ad}} \qquad (1.239)$$

for the average pressure perturbation $\delta\bar{p}$, we have

$$\delta\bar{p}=\bar{\Gamma}_1\bar{p}\,\frac{\delta\bar{\varepsilon}}{\bar{\varepsilon}}. \qquad (1.240)$$

Substituting relation (1.240) into the averaged perturbation equation (1.236) and taking into account equations (1.234) and (1.235), we can obtain

$$\delta\ddot{a}+\omega^2\delta a=0, \qquad (1.241)$$

where the following notation is introduced:

$$\omega^2=\frac{M}{a^3}\{3\,(\bar{\Gamma}_1-1)\,(1+M/a)\,(1-3M/a)\,(1-3M/a$$
$$+6M^2/a^2)-(1-6M/a+3M^2/a^2)\,(1+2M/a+O\,(M^2/a^2)\}$$
$$\times\,[1+2M/a+O\,(M^2/a^2)]^{-2}\,[1-2M/a+3M^2/a^2]^{-1}. \qquad (1.242)$$

When $\omega^2>0$, this equation describes the harmonic oscillation of the astronomical object's radius about an equilibrium position and,

therefore, when $\omega^2 > 0$ the astronomical object will be stable with respect to small perturbations of its radius which do not change the total rest mass of its constituent particles.

When $\omega^2 < 0$, equation (1.241) describes a temporary exponential increase in the initial perturbation δa (0). Thus. when $\omega^2 \leqslant 0$, a static, spherically symmetric object will be unstable with respect to small perturbations of its radius.

Let us determine the values of the mean adiabatic exponent $\overline{\Gamma}_1$ of matter, at which the astronomical object will be unstable. Since

$$(1 - 2M/a + 3M^2/a^2) > 0,$$

the condition $\omega^2 \leqslant 0$ and expression (1.242) suggest that a static, spherically symmetric object will be unstable to small perturbations of its radius, once the mean adiabatic exponent satisfies the following inequality:

$$\overline{\Gamma}_1 \leqslant \frac{[4 - 19M/a + O(M^2/a^2)]}{3[1 - 3M/a + 6M^2/a^2][1 - 3M/a][1 + M/a]} = F\left(\frac{M}{a}\right). \tag{1.243}$$

To be rigorous, this formula is applicable only within the range of the averaged gravitational parameter $M/a < 1$, since we are using the approximate expansion (1.234). Yet we can extrapolate this expression in a qualitative fashion over the range $M/a \sim 1$ as well. The function $F(x)$ at $x = 0$ equals 4/3 and as $x \to 1/3$, it diminishes to minus infinity. Thus, once the value of the mean adiabatic exponent of astronomical object matter turns out to be larger than a value due to (1.239) in a certain range of the M/a values, an astronomical object will be stable to small perturbations of its radius within this same range.

It appears from the result of observations [106], that main sequence stars have a fairly low temperature, hence gas pressure in these stars dominates over radiation pressure, and this results in the adiabatic exponent being equal to $\overline{\Gamma}_1 = 5/3$. The radiation pressure may be higher than gas pressure for higher temperature stars and here the adiabatic exponent value is slightly greater than 4/3, i.e. $\overline{\Gamma}_1 \gtrsim 4/3$. However, stars [106] such as these have not yet been observed, since their lifetime must be rather short.

Yet it should be noted that the aforementioned values of the adiabatic exponent may only be used if the averaged pressure inside the star is positive, i.e. $\overline{p} > 0$. The expression for the averaged pressure can easily be obtained from (1.235):

$$\overline{p} = \frac{\overline{\varepsilon} M(1 - 3M/a)/a}{1 - 3M/a + 6M^2/a^2}. \tag{1.244}$$

This expression suggests that when $M/a < 1/3$, the averaged pressure is positive and negative when $M/a > 1/3$.

Thus, our qualitative analysis makes it possible to state that when $M/a < 1/3$ the averaged potential• values of astronomical objects will be stable for small radial perturbations, and when $M/a \geqslant 1/3$, unstable.

Thus the passage from stability to instability will be accomplished jumpwise, i.e. from an infinitely high stability to an infinitely high instability. This means that the instability could be explosive, so that any small perturbation of the astronomical object that violates its equilibrium could bring about a matter ejection.

Consider a spherically symmetric, static astronomical object. What would be the result of an accretion of matter by this object from its environment? It follows from expression (1.237) that a change in the object's mass would inevitably be accompanied by a change in the mean density of its matter and in its radius:

$$\frac{\delta M}{M} = \frac{\delta \bar{\varepsilon}}{\bar{\varepsilon}} [1 + 3M/a + O(M^2/a^2)] + 3\frac{\delta a}{a} \times [1 + 2M/a + O(M^2/a^2)]. \tag{1.245}$$

Yet changing the mean density of the matter of the astronomical object and its radius would naturally give rise to a change in its mean gravitational potential:

$$\delta\left(\frac{M}{a}\right) = \frac{\delta \bar{\varepsilon}}{\bar{\varepsilon}} \frac{M}{a} [1 + 3M/a + O(M^2/a^2)] + 2\frac{\delta a}{a}\frac{M}{a} [1 + 3M/a + O(M^2/a^2)]. \tag{1.246}$$

Let us determine the change in the value of the mean gravitational potential of the astronomical object which accretes mass in the two following limiting cases, each corresponding to a change in only one of the object's two parameters (its mean density or its radius).

Assume that the averaged density of the astronomical object's matter does not change during the mass accretion, i.e. $\delta\bar{\varepsilon} = 0$. According to Eq. (1.245), the mass accretion δM will then result in changing the object's radius as follows:

$$\frac{\delta a}{a} = \frac{\delta M}{3M} [1 - 2M/a + O(M^2/a^2)]. \tag{1.247}$$

Substituting this expression into identity (1.246) yields

$$\delta\left(\frac{M}{a}\right) = \frac{2}{3}\frac{\delta M}{a} [1 + M/a + O(M^2/a^2)]. \tag{1.248}$$

Thus, if the mean density of the matter is unchanged, the mass accretion by the astronomical object will bring about an increase in its mean gravitational potential.

Suppose the object's radius is unchanged when accreting the mass, i.e. $\delta a = 0$. The mass accretion δM will then lead to a change in the mean density of the object's matter:

$$\frac{\delta\bar{\varepsilon}}{\bar{\varepsilon}} = \frac{\delta M}{M}\,[1 - 3M/a + O\,(M^2/a^2)]. \tag{1.249}$$

We can obtain the change in the mean gravitational potential of the object from expressions (1.246) and (1.249):

$$\delta\left(\frac{M}{a}\right) = \frac{\delta M}{a}\,[1 + O\,(M^2/a^2)]. \tag{1.250}$$

As a consequence, in this case too, the mass accretion by the astronomical object will lead to an increase in its mean gravitational potential.

Thus, any astronomical object tends to have its mean gravitational potential increased due to the gravitational accretion of the surrounding matter.

The above qualitative analysis of the evolution of astronomical objects demonstrates that in the field theory of gravitation with minimum coupling objects in the range of the mean gravitational potential $M/a < 1/3$ are stable to small perturbations of the object's radius, the rest mass being unchanged.

The value of the mean gravitational potential of these objects, however, increases as they accrete surrounding matter. When the averaged potential $M/a = 1/3$, the object jumps from being stable to being unstable with respect to small radial perturbations. Therefore, once the critical value of the mean gravitational potential is attained, even small radial perturbations will lead to an expansion of the matter that may be accompanied by an ejection of a certain part of the object's mass and an energy release.

Consequently, a new mechanism for energy release exists in the given theory in place of the gravitational collapse that takes place in the GTR for unstable astronomical objects.

1.16 THE GRAVITATIONAL FIELD OF A NON-STATIC, SPHERICALLY SYMMETRIC SOURCE

According to the Birkhoff theorem, the gravitational field of a non-static, spherically symmetric source in Einstein's theory is a static field outside the matter with a metric corresponding to the Schwarzschild solution.

We will show that in the field theory of gravitation the gravitational field outside matter is also a static field in the case of a non-static, spherically symmetric source and has components expressed by formulas (1.219) and (1.220). Consider the case of matter distributed inside a sphere of radius a in a spherically symmetric fashion and moving radially in the same fashion.

From the symmetry of the problem the tensor components T^{ni}, h^{ni}, I_{ni}, and f_{ni} that will not vanish will be the diagonal components and also the components T^{0r} , I_{0r}, h^{0r} and f_{0r}. All the components of these tensors, except the ($\varphi\varphi$) ones, will depend on r and t. For the ($\varphi\varphi$)-components we have

$$T^{\varphi\varphi}=\frac{T^{\theta\theta}}{\sin^2\theta}, \quad h^{\varphi\varphi}=\frac{h^{\theta\theta}}{\sin^2\theta},$$
$$I_{\varphi\varphi}=I_{\theta\theta}\sin^2\theta, \quad f_{\varphi\varphi}=f_{\theta\theta}\sin^2\theta.$$

The four-vector for the velocity of the matter is of the form

$$u^i=\{u^0(r,\ t),\ u^r(r,\ t),\ 0,\ 0\}.$$

Expanding the tensor current I_{lm} and the gravitational field f_{lm} components into the Fourier integrals in the time variable, we get

$$f_{lm}=\int d\omega \exp(-i\omega t)\, f_{lm}(\omega,\ r),$$
$$I_{lm}=\int d\omega \exp(-i\omega t)]I_{lm}(\omega,\ r).$$

Suppose we single out the static part $I_{lm}(r)$ of the spectrum $I_{lm}(\omega,\ r)$. It will obviously yield the static solutions treated in the foregoing section. Therefore, when $I_{ml}(\omega,\ r)$ is used below the non-static part will be implied.

The field equations (1.91) for the cases under consideration will be the following ordinary differential equations:

$$\left.\begin{aligned} & f''_{00}+\frac{2}{r}f'_{00}+\omega^2 f_{00}=16\pi I_{00}, \\ & f''_{0r}+\frac{2}{r}f'_{0r}+\left(\omega^2-\frac{2}{r^2}\right)f_{0r}=16\pi I_{0r}, \\ & f''_{rr}+\frac{2}{r}f'_{rr}+\frac{4}{r^2}f^{\theta}_{\theta}+\left(\omega^2-\frac{4}{r^2}\right)f_{rr}=16\pi I_{rr}, \\ & f^{\theta''}_{\theta}+\frac{2}{r}f^{\theta'}_{\theta}-\frac{2}{r^2}f_{rr}+\left(\omega^2-\frac{2}{r^2}\right)f^{\theta}_{\theta}=16\pi I^{\theta}_{\theta}. \end{aligned}\right\} \quad (1.251)$$

It is only natural to require that the functions f_{00}, f_{0r}, f^{θ}_{θ} and f_{rr} be bounded as $r\to 0$ and that the radiation conditions are satisfied as $r\to\infty$ and that these be the boundary conditions for these equations. Since the tensor current conserved, $D_l I^l_n = 0$, we have

$$\left.\begin{aligned} & i\omega I_{00}+I'_{0r}+(2/r)\,I_{0r}=0, \\ & i\omega I_{[illegible]}+I'_{[illegible]}+(2/r)\,I_{[illegible]}-(2/r^3)\,I_{[illegible]}=0 \end{aligned}\right\} \quad (1.252)$$

Solving equations (1.251) and allowing for (1.252), we shall obtain

$$f_{rr} = (A_1 + 2A_2)/3;$$
$$f_{\theta\theta} = (r^2/3)(A_1 - A_2);$$
$$f_{00} = -\frac{8\pi^2}{\sqrt{r}}\Big\{H^{(1)}_{1/2}(\omega r)\int_0^r x^{3/2}\,dx\,I_{0r}J_{3/2}(\omega x)$$
$$+ J_{1/2}(\omega r)\int_r^a x^{3/2}\,dx\,I_{0r}H^{(1)}_{3/2}(\omega x)\Big\},$$
$$f_{0r} = -\frac{8\pi^2 i}{\sqrt{r}}\Big\{H^{(1)}_{3/2}(\omega r)\int_0^r x^{3/2}\,dx\,J_{3/2}(\omega x)\,I_{0r}$$
$$+ J_{3/2}(\omega r)\int_r^a x^{3/2}\,dx\,H^{(1)}_{3/2}(\omega x)\,I_{0r}\Big\},$$

where the following notations were introduced:

$$A_1 = -\frac{8\pi^2 i\omega}{\sqrt{r}}\Big\{H^{(1)}_{1/2}(\omega r)\int_0^r x^{5/2}\,[iI_{0r}J_{1/2}(\omega x)$$
$$+ I_{rr}J_{3/2}(\omega x)]\,dx + J_{1/2}(\omega r)\int_r^a x^{5/2}\,[iI_{0r}H^{(1)}_{1/2}(\omega x)$$
$$+ I_{rr}H^{(1)}_{3/2}(\omega x)]\,dx\Big\},$$
$$A_2 = \frac{4\pi^2 i\omega}{\sqrt{r}}\Big\{H^{(1)}_{5/2}(\omega r)\int_0^r x^{5/2}\,dx\,[iI_{0r}J_{5/2}(\omega x)$$
$$- I_{rr}J_{3/2}(\omega x)] + J_{5/2}(\omega r)\int_r^a x^{5/2}\,dx\,[iI_{0r}H^{(1)}_{5/2}(\omega x)$$
$$- I_{rr}H^{(1)}_{3/2}(\omega x)]\Big\}.$$

Using the gauge transformation

$$f_{ni} \to f_{ni} + D_n a_i + D_i a_n - \gamma_{ni} D_l a^l$$

outside the matter, we can impose two conditions on the gravitational field components:

$$f = 0, \quad f^{00} = 0.$$

The gauge four-vector has to satisfy the equation $D_l D^l a_i = 0$ outside matter, lest the gauge transformation violates the condition $D_l f^{lm} = 0$. By choosing the gauge vectors in the form

$$a_0 = \frac{2\pi^2 i}{\omega \sqrt{r}} H^{(1)}_{1/2}(\omega r) \int_0^a x^{3/2}\, dx \,\{I_{0r}[J_{3/2}(\omega x) + \omega x J_{1/2}(\omega x)] - i\omega x I_{rr} J_{3/2}(\omega x)\};$$

$$a_r = \frac{2\pi^2}{\sqrt{r}} H^{(1)}_{3/2}(\omega r) \int_0^a x^{5/2}\, dx\, [I_{0r} J_{5/2}(\omega x) + i I_{rr} J_{3/2}(\omega x)],$$

$$a_\theta = a_\varphi = 0,$$

one can be easily sure that all the non-static gravitational field components will vanish outside matter:

$$f_{nm} = 0.$$

Thus, in the case of a non-static source, for which the distribution and motion of the matter is spherically symmetric, the gravitational field outside the matter will be static with its components defined by formulas (1.219) and (1.220).

1.17 A NON-STATIONARY MODEL OF THE HOMOGENEOUS UNIVERSE

The field theory of gravitation enables non-stationary models of the Universe to be constructed that can describe the cosmological red shift and yet are free from Newtonian-type divergences. These models fit a flat Universe.

It should be noted that a model of the Universe in the field theory of gravitation characterizes only that part of it which has a linear dimension $r \sim cT$, T being the Universe age. From this point of view the "birth" of the Universe means that in the past the matter density was sufficiently great within a given, sufficiently large domain of the Universe. The subsequent evolution of this domain may be described by the model under consideration. Any other domain of the Universe may thereby develop independently of the development of the given domain and perhaps according to absolutely different laws. Yet observing them is possible in the field theory of gravitation.

Astronomical observations [107, 108] show a quite heterogeneous distribution of matter within the Universe, the mass being mainly concentrated in the planets and stars, with interstellar gas and radiation comprising only a small fraction of the total mass. However, averaging over space domains with linear dimensions much greater than the distance between the aggregates of galaxies, the

density of the matter for the part of the Universe accessible to observation appears to be constant and independent of the center of the averaging domain. Hence, from the standpoint of physics it would be natural to consider, as a first step, a model of an homogeneous isotropic Universe.

With this approach, the heterogeneity appears when averaging over smaller space domains (galaxies aggregates, galaxies themselves, etc.) can be allowed for by introducing small heterogeneous perturbations into the background cosmological field of a homogeneous Universe. The homogeneous isotropic Universe is described by an interval

$$ds^2 = U(t)\,dt^2 - V(t)\,[dx^2 + dy^2 + dz^2]. \tag{1.253}$$

The matter in the Universe will be treated as a perfect liquid having the energy-momentum tensor density

$$T^{ni} = \sqrt{-g}\,[(\mathscr{E} + p)\,u^i u^n - p g^{ni}].$$

By virtue of homogeneity and isotropy of the Universe, we have

$$\mathscr{E} = \varepsilon(t), \quad p = p(t), \quad u^\alpha = 0, \quad u^0 \neq 0, \quad u^0 u^0 g_{00} = 1.$$

The energy-momentum tensor density components for matter will now take the form

$$T^{00} = \varepsilon \sqrt{\frac{V^3}{U}}, \qquad T^{\alpha\beta} = -p\sqrt{UV}\,\gamma^{\alpha\beta}. \tag{1.254}$$

Using expression (1.253) for the interval, let us determine the Riemannian space-time connections

$$\begin{aligned} &\Gamma^0_{00} = \frac{\dot U}{2U}, \qquad \Gamma^0_{0\alpha} = 0, \quad \Gamma^\alpha_{00} = 0, \\ &\Gamma^0_{\alpha\beta} = \frac{\dot V}{2U}\,\gamma_{\alpha\beta}, \quad \Gamma^\alpha_{0\beta} = \frac{\dot V}{2V}\,\delta^\alpha_\beta, \quad \Gamma^\alpha_{\beta\varepsilon} = 0, \end{aligned} \tag{1.255}$$

the dot standing for ordinary differentiation with respect to time.

Substituting expressions (1.254) and (1.255) into the covariant conservation equation of the energy-momentum tensor density for matter (1.198), we obtain

$$\frac{d}{dt}\left(\varepsilon\sqrt{V^3}\right) + p\,\frac{d}{dt}\sqrt{V^3} = 0. \tag{1.256}$$

The solution of equation (1.256) has the form

$$\ln V = -\frac{2}{3}\int\limits_{\varepsilon_0}^{\varepsilon} \frac{d\varepsilon'}{\varepsilon' + p(\varepsilon')}. \tag{1.257}$$

A coupling equation can be written in the most general form as

$$g_{ni} = \gamma_{ni} f_1 + f_{ni} f_2 + f_{im} f_{nl} A^{lm}, \tag{1.258}$$

where f_1 and f_2 are certain scalar functions of the invariants $I_1 = f$, $I_2 = f_{nm}f^{nm}$, etc., and the tensor A^{lm} is constructed from the tensors γ^{nm}, f^{nm}, $f^{ni}f_i^m$, . . . and from the invariants.

The gravitational field equations (1.91) for an homogeneous Universe will take the form

$$\ddot{f}_{00} = \ddot{f}_{0\alpha} = 0, \quad f_{\alpha\beta} = -16\pi\{h_{\alpha\beta} + \gamma_{\alpha\beta}h_{00}\}. \tag{1.259}$$

Using definition (1.70), we obtain

$$f_{00} = 0, \quad f_{0\alpha} = 0. \tag{1.260}$$

Due to the isotropy of the Universe, the spatial components of gravitational field have to be of the form

$$f_{\alpha\beta} = \gamma_{\alpha\beta}F(t). \tag{1.261}$$

Then

$$g_{00} = U(F), \quad g_{\alpha\beta} = \gamma_{\alpha\beta}V(F). \tag{1.262}$$

It can be shown that

$$\frac{\partial g_{00}}{\partial f_{\alpha\beta}} = \frac{1}{3}\gamma^{\alpha\beta}\frac{dU}{dF}, \quad \gamma^{\varepsilon\eta}\frac{\partial g_{\varepsilon\eta}}{\partial f_{\alpha\beta}} = \gamma^{\alpha\beta}\frac{dV}{dF}.$$

Hence field equations (1.259) will take the form

$$\ddot{F} = \frac{64\pi}{3}\left\{\varepsilon\sqrt{V^3}\frac{d}{dF}\sqrt{U} - p\sqrt{U}\frac{d}{dF}\sqrt{V^3}\right\}. \tag{1.263}$$

The conditions currently existing at $t = 0$ will be taken as the initial conditions for equation (1.263), viz.

$$\varepsilon = \varepsilon_0, \quad U = V = 1, \quad \frac{dV}{dt} = 2H, \tag{1.264}$$

H being the Hubble constant. It should be specially noted that the initial conditions are chosen having made the assumption that the energy density of matter $\varepsilon_0 \neq 0$. Therefore, the subsequent analysis too will pertain to only this case. It follows from experience [17] that

$$20\cdot 10^9 \text{ years} > \frac{1}{H} > 7.5\cdot 10^9 \text{ years}.$$

With the initial conditions chosen in such a fashion the cosmological field will be at the present instant, the pseudo-Euclidean background against which every other physical process is considered.

It follows from conditions (1.264) that

$$F(0) = 0, \quad \left.\frac{dF}{dt}\right|_{t=0} = -4H.$$

Equation (1.263) reduces to

$$\frac{d}{dt}[\dot{F}^2 + C_1] = \frac{128\pi}{3}\left\{\varepsilon\sqrt{V^3}\frac{d}{dt}\sqrt{U} - p\sqrt{U}\frac{d}{dt}\sqrt{V^3}\right\}.$$

Taking into account conservation equation (1.256), we obtain

$$\dot{F}^2+C_1=\frac{128\pi}{3}\varepsilon\sqrt{V^3U}. \tag{1.265}$$

It is of interest to note that equation (1.265) is a modified way of writing the energy density conservation law of matter and gravitational field in a flat space-time. Actually, once definitions (1.105) and (1.106), the coupling equation (1.258), and the components of the energy-momentum tensor density of matter in Riemannian space-time (1.254) are used and identities (1.260) and (1.261) are taken into account, we will obtain

$$t_{\mathrm{M}}^{00}=\varepsilon\sqrt{UV^3}, \quad t_{\mathrm{g}}^{00}=-\frac{3}{128\pi}\dot{F}^2, \quad t_{\mathrm{M}}^{0\alpha}=t_{\mathrm{g}}^{0\alpha}=0. \tag{1.266}$$

Therefore, the conservation law of the energy-momentum tensor density in flat space-time (1.97) will have the form

$$\frac{\partial}{\partial t}[t_{\mathrm{M}}^{00}+t_{\mathrm{g}}^{00}]=0.$$

Whence it follows that $t_{\mathrm{M}}^{00}+t_{\mathrm{g}}^{00}=\mathrm{const}$.

By using initial conditions (1.264) and expression (1.266), we can obtain

$$\varepsilon\sqrt{V^3U}-\frac{3}{128\pi}\dot{F}^2=\frac{3}{128\pi}C_1,$$

where

$$C_1=16H^2(\alpha-1), \quad \alpha=\frac{8\pi\varepsilon_0}{3H^2}.$$

Thus, the total energy density of the matter and the gravitational field of the Universe is constant at every stage of its evolution in flat space-time. This reveals that the energy of the Universe does not change in the course of its evolution, being only redistributed between matter and the gravitational field.

Using the initial conditions, the solution of equation (1.265) can be written as

$$t=-\frac{1}{4H}\int_0^F\frac{dF'}{\sqrt{1-\alpha+\frac{\alpha\varepsilon}{\varepsilon_0}\sqrt{V^3U}}}. \tag{1.267}$$

Expressions (1.257), (1.262), and (1.267) parametrize the whole evolution of a homogeneous, isotropic Universe including the singular state (or the hot Universe) for an arbitrary equation of state for the matter, $p=p(\varepsilon)$, and coupling equation (1.258) written in the most general form.

In transforming to proper time in expressions (1.262) and (1.253), it is possible to transform to a proper time $\tau(t)$ within a time interval which has a non-vanishing $U(t)$, such that

$$\sqrt{U(t)}\,dt = d\tau.$$

The interval will then look like

$$ds^2 = d\tau^2 - V(\tau)\,[dx^2 + dy^2 + dz^2]. \tag{1.268}$$

Assuming that the present instant $\tau(0) = 0$, we obtain parametric expressions defining the evolution of the Universe:

$$\tau = -\frac{1}{4H}\int\limits_0^F \frac{\sqrt{U}\,dF'}{\sqrt{1-\alpha+(\alpha\varepsilon/\varepsilon_0)\sqrt{UV^3}}}, \tag{1.269}$$

$$\ln V(F) = -\frac{2}{3}\int\limits_{\varepsilon_0}^{\varepsilon} \frac{d\varepsilon'}{\varepsilon'+p(\varepsilon')}. \tag{1.270}$$

From the minimum-coupling equation (1.95) we have the following for the functions U and V:

$$\left.\begin{aligned} U &= 1-\frac{3}{2}F+\frac{1}{4}(9b_4+3b_3)F^2, \\ V &= 1-\frac{1}{2}F+\frac{1}{4}(b_1+3b_2+3b_3+9b_4)F^2. \end{aligned}\right\} \tag{1.271}$$

Let us study the solutions obtained for the vicinity of the present instant $\left(|\tau| \ll \frac{1}{4H}\right)$ in proper time. We shall assume that the pressure in the vicinity of the present instant of the proper time is negligibly small compared to the energy density, $p \ll \varepsilon$. Hence we can obtain from equation (1.270) that

$$\varepsilon = \frac{\varepsilon_0}{\sqrt{V^3(F)}}. \tag{1.272}$$

Substituting the expressions for U, V, and ε into integral (1.269) and performing integration, we have

$$\tau = -\frac{1}{4H}\left[F-\frac{3}{8}\left(1-\frac{\alpha}{2}\right)F^2+O(F^3)\right].$$

Solving for F in this relationship and substituting it into the expression for $V(F)$, we get

$$V(\tau) = 1+2H\tau^2+H^2\tau^2\left[\frac{3}{2}\alpha - 3+4(b_1+3b_2+3b_3+9b_4)\right]+O(H^3\tau^3).$$

The metric (1.268) with the cosmological scale factor $V(\tau)$ results in experimentally observable effects. One such effect is the

cosmological red shift discovered by Hubble in 1929 [109]. It is the shift to be reddened of the spectral lines emitted by remote galaxies and its magnitude is proportional to the distance separating the galaxies and the Earth. This effect was obtained from the GTR by the Soviet scientist A. A. Friedman [110] in 1922.

In the field theory of gravitation the model of a homogeneous Universe also describes the cosmological red shift of frequency in the vicinity of the present (at $H\tau \ll 1$ or $\tau \ll 10^{10}$ years):

$$\Delta\omega = -HL\omega.$$

The deceleration parameter of an expanding Universe, $q = 1 - 2V\ddot{V}/\dot{V}^2$, is equal to

$$q_0 = 4 - \frac{3}{2}\alpha - 4(b_1 + 3b_2 + 3b_3 + 9b_4) \tag{1.273}$$

in the vicinity of the present, $\tau = 0$.

By comparison in Einstein's theory the deceleration parameter of the homogeneous Universe is equal to

$$q_0 = \frac{\alpha}{2}.$$

The deceleration parameter in Einstein's theory of gravitation is one of the most important quantities for the homogeneous Universe as a whole, since if the deceleration parameter value is $q_0 < 1/2$ $(\alpha < 1)$ the Universe is open, and when $q_0 > 1/2$ $(\alpha > 1)$ the Universe is closed and has a finite volume with no boundaries. The field theory of gravitation lacks this interplay—the Universe has an infinite volume for any value of α and q_0.

The estimated matter density of the galaxies [107] is

$$\varepsilon_0 = 3 \cdot 10^{-31} \text{ g/cm}^3.$$

Thus

$$\alpha = 0.06, \tag{1.274}$$

and the deceleration parameter has to be $q_0 = 0.03$ in Einstein's theory which results in an open Universe with an infinite expansion. However, measurements of the deceleration parameter have produced the different result.

As an example, [111] concludes that the value of q_0 lies in the range from 2 to 32, the most probable value being $q_0 = 5$. Thus, the value of the deceleration parameter obtained from observation contradicts in Einstein's theory with the matter density observed within the galaxies, which is much smaller than that required to match. To eliminate the discrepancy between the characteristics of the cosmological version of Einstein's theory and the values obtained from observations, attempts have started to be made both to increase the value of ε_0 (the search for the deficient matter in the galaxies,

the "hidden mass mystery") and to decrease the value of q_0 obtained from experiment (an assumption of the pronounced evolution of the luminosity function depending on the red shift value). These attempts have not, to date, brought any certainty to the question.

In the field theory of gravitation, as opposed to Einstein's GTR, not only is the deceleration parameter governed by the mean density of matter ε_0 (the parameter $\alpha = 8\pi\varepsilon_0/3H^2$), but also by the minimum coupling parameters. Therefore, measuring the deceleration parameter q_0 enables the following quantity to be measured without referring to the post-Newtonian experiments in the solar system:

$$b_1 + 3b_2 + 3b_3 + 9b_4 = -\frac{1}{4}\left[q_0 + \frac{3}{2}\alpha - 4\right]. \tag{1.275}$$

The behavioral pattern of the model of the homogeneous Universe in the remote past is substantially dependent on the form of the coupling equation in strong gravitational fields.

Provided that equation $V(F_1) = 0$ has real roots, the determinant of the metric tensor and its spatial components will vanish at $F = F_1$. For this reason it is only natural to assume that a singular state of the Universe is realized at $F = F_1$. Ultrarelativistic particles dominate near the singular state, equation of state being of the form

$$p = \varepsilon/3.$$

Putting this equation into expression (1.270), we obtain

$$\varepsilon = \varepsilon_0/V^2. \tag{1.276}$$

Whence it follows that, once the function $V(F)$ vanishes, the total energy density of the Universe goes to infinity and the singular state of the Universe is realized when $F = F_1$.

Some instant in past, $\tau = \tau_m$, corresponds to the least positive root F^* of the equation $V(F) = 0$. The time $T = -\tau_m$ may naturally be called the Universe age, viz.

$$T = \frac{1}{4H}\int_0^{F^*} \frac{\sqrt{U}\,dF}{\sqrt{1-\alpha+\alpha(\varepsilon/\varepsilon_0)\sqrt{UV^3}}}. \tag{1.277}$$

Let us introduce the time $\tau_0 = T + \tau$ reckoned from the singular state:

$$\tau_0 = \frac{1}{4H}\int_F^{F^*} \frac{\sqrt{U}\,dF}{\sqrt{1-\alpha+\alpha(\varepsilon/\varepsilon_0)\sqrt{UV^3}}}.$$

Relation (1.276) is valid in the vicinity of the singular state (at $F \sim F^*$), hence we obtain

$$\tau_0 = \frac{1}{4H}\int_F^{F^*} \frac{\sqrt{U}\,dF}{\sqrt{1-\alpha+\alpha\sqrt{U/V}}}. \tag{1.278}$$

Expression (1.278) defines the proper time dependence in the vicinity of the singular state on the gravitational field F and thus enables the behavior of the function $V(\tau)$ to be specified within the given neighborhood.

Note that if the equation $V(F_1) = 0$ has no real roots, the model of the Universe has no singular state. In this situation Olbers paradox may emerge, i.e. a divergency in the luminosity integral for all stars. Actually, at the present instant $\tau = 0$, the total energy ρ of the stellar light is equal [17] to

$$\rho = \int_{-\infty}^{0} Z(\tau) V^2(\tau)\, d\tau, \tag{1.279}$$

where $Z(\tau)$ is the proper luminosity density for stars:

$$Z(\tau) = \int n(\tau, L)\, dL,$$

$n(\tau, L)$ being the density of stars with an absolute luminosity L at an instant τ. For the integral (1.279) to converge, either a singular state of the Universe ($V(F_1) = 0$) must exist in finite F_1, resulting in the effective cut-off of integral (1.279) at the lower limit, some $\tau = \tau(F_1)$, or the quantity $V(\tau)$, sufficiently fast, must vanish with increasing $|\tau|$, viz.

$$\tau V(\tau) Z(\tau) \to 0, \quad |\tau| \to \infty. \tag{1.280}$$

Let us introduce the following notations:

$$\begin{aligned} &b_1 + 3b_2 + 3b_3 + 9b_4 = w, \\ &3b_3 + 9b_4 = k. \end{aligned} \tag{1.281}$$

Incorporating these notations, expression (1.271) can be rewritten as

$$\begin{aligned} U &= 1 - \frac{3}{2} F + \frac{k}{4} F^2, \\ V &= 1 - \frac{1}{2} F + \frac{w}{4} F^2. \end{aligned} \tag{1.282}$$

Let us study how the coefficients k and w influence the behavior pattern of the model of the Universe. We will require that the theory is devoid of both an Olbers type paradox and physical singularities of the Universe metric at the finite values of matter and energy density. As follows from expression (1.282), the first of these requirments can be met only when

$$w \leqslant 1/4. \tag{1.283}$$

By virtue of relations (1.273) and (1.281), this condition implies that the Universe deceleration parameter is limited in the vicinity of the present:

$$q_0 \geqslant 3 - \frac{3}{2}\alpha. \tag{1.284}$$

The second requirement imposes limitations on the values of the real roots of the equation $U(F) = 0$ and, thereby, on the value of the coefficient k.

Depending on the values of k and w, various types of model of the Universe are feasible. Consider them one-by-one.

I. $$0 \leqslant w \leqslant 1/4 \text{ or } 4 - (3/2)\,\alpha \geqslant q_0 \geqslant 3 - (3/2)\,\alpha. \tag{1.285}$$

Here both roots of the function V are positive. The least one, pertinent to a singular state of the Universe, is

$$F^* = \frac{1 - \sqrt{1 - 4w}}{w}.$$

The values of the root F^* over the above range of w (1.285) are confined within

$$2 \leqslant F^* \leqslant 4.$$

Since negative F values correspond to future evolution of the Universe, for the case of (1.285) the Universe will "expand" infinitely long. Its metric (1.282) will thus have no singularities outside the Universe singular state, provided that the function U does not vanish within the range $-\infty < F < F^*$. It can easily be shown that this is only possible when the condition

$$k > 9/4 \tag{1.286}$$

is satisfied.

The value of the temporal factor U when the Universe is in the singular state is

$$U(F^*) = -F^* + \frac{k - w}{4} F^{*2}. \tag{1.287}$$

Whence it follows that in range (1.285) for w, the value of the temporal factor U at $F = F^*$ satisfies the inequalities

$$-5 + 4k \geqslant U \geqslant -2 + k > 1/4.$$

The nature of the Universe evolution in the vicinity of the singular state is essentially influenced by the value of the parameter w. For example, if $w = 1/4$, we have from expressions (1.278), (1.276), and (1.282)

$$V \sim \tau_0^{4/3}, \ \varepsilon \sim \tau_0^{-8/3}, \ U \simeq -5 + 4k.$$

If $w = 0$, we obtain

$$V \sim \tau_0^{4/5},\ \varepsilon \sim \tau_0^{-8/5},\ U \simeq -2+k.$$

Thus, at large values of w, the Universe singular state grows faster in the course of time. To compare, the estimates in the GTR [11] are

$$V \sim \tau_0,\ \varepsilon \sim \tau_0^{-2}$$

and are valid in the vicinity of the singular state for any model of the Universe.

The behavior pattern of the functions U and V near the singularity substantially controls the flux densities and the spectral characteristics of remnants of electromagnetic, neutrino, and gravitational radiations. The frequency of the electromagnetic and neutrino radiation and, consequently, their temperature change as a result of both the influence of the cosmological gravitational field and the Doppler effect. Yet the frequency and the temperature of the gravitational radiation can only be changed by the Doppler effect. Therefore, measuring the flux densities and spectral characteristics of those remnants of radiations enables the behavior of the coupling equation to be specified in strong gravitational fields.

To find the Universe age requires an equation of state for the matter. Inasmuch as the exact equation of matter state is unknown, we can only evaluate the Universe age in an approximate fashion. Note, first of all, that with any equation of matter state the inequality

$$0 < p \leqslant \varepsilon/3$$

holds. By virtue of (1.270) and (1.282) this means that when $0 \leqslant F \leqslant F^*$, the estimate

$$\frac{\varepsilon_0}{\sqrt{V^3}} \leqslant \varepsilon \leqslant \frac{\varepsilon_0}{V^2}$$

is valid. Hence we have for the Universe age

$$T_1 \geqslant T \geqslant T_2,$$

where

$$T_1 = \frac{1}{4H} \int_0^{F^*} \frac{\sqrt{U}\, dF}{\sqrt{1-\alpha+\alpha\sqrt{U}}},$$

$$T_2 = \frac{1}{4H} \int_0^{F^*} \frac{\sqrt{U}\, dF}{\sqrt{1-\alpha+\alpha\sqrt{U/V}}}.$$

An analysis will show that the numerical values of the quantities T_1 and T_2 are essentially dependent on the values of the parameters w and k, which can change T_1 and T_2 considerably. A minimum value of

the Universe age within the range of the parameters (1.285) and (1.286) is achieved when $w=0$ and $k=9/4$, viz.

$$\frac{3}{4H} \geqslant T \geqslant \frac{2}{9H}.$$

Increasing the parameters w and k within the range of (1.285) and (1.286) increases the Universe age monotonically. For example, at $w = 1/4$, $k = 12$, we have

$$\frac{7}{H} \geqslant T \geqslant \frac{1}{H}.$$

It should be stressed that for the case under consideration, the minimum-coupling parameters b_1 and b_2, as well as b_3 and b_4, do not vanish in pairs due to inequalities (1.221), (1.285), and (1.286). Moreover, given that one of the parameters b_1, b_2, b_3, or b_4 vanishes, all the remaining parameters are necessarily nonzero.

II. $\quad w < 0$ or $q_0 > 4 - (3/2)\,\alpha.$

Here the roots of the function V have different signs and, as a consequence, an "expansion" of the Universe will change into a "contraction" in the future and it will return to the singular state. This change-over will take place when the scale factor V becomes

$$V = 1 - 1/(4w) > 1.$$

The root F^* corresponds to the initial state of the Universe and is confined to

$$0 < F^* < 2.$$

The return of the Universe to the singular state will take place at

$$F_2 = \frac{1+\sqrt{1-4w}}{w}.$$

It can easily be checked that the value of F_2 is confined within the limits

$$0 > F_2 > -\infty.$$

Note that the Universe metric will have singularities between those singular states unless the function U has no roots within the range

$$F_2 < F < F^*.$$

The character of the Universe evolution in the vicinity of the singular state as well as the Universe age will be substantially influenced by the values of the parameters w and k. The Universe age can, as an analysis demonstrates, be both more and less than $1/H$.

Thus, non-stationary homogeneous models of the Universe in the field theory of gravitation deal with the cosmological red shift and allow for both monotonic and non-monotonic behavior. The behavior pattern of the model and the Universe age are substantially dependent on the value of the deceleration parameter q_0, i.e. when q_0 is confined within the limits $4 - (3/2)\,\alpha \geqslant q_0 \geqslant 3 - (3/2)\,\alpha$, the Universe will "expand" for an infinitely long time, and when $q_0 > 4 - (3/2)\,\alpha$, the "expansion" will give way to a "contraction" in the course of time so that the Universe will return to the singular state.

1.18 FEASIBLE EXPERIMENTS IN THE SEARCH FOR DIFFERENCES BETWEEN THE PREDICTIONS OF THE FIELD THEORY OF GRAVITATION AND THOSE OF EINSTEIN'S GTR

The field theory of gravitation and Einstein's GTR are absolutely different theories of gravitation, since the fundamentals of these theories and the gravitational field equations are different. Consequently, these theories will yield the dissimilar predictions for the same physical situation.

The difference between them should be especially clear-cut when describing gravitational waves and the effects due to strong gravitational fields. It should be stressed that since the Einsteinian quadrupole formula is not a part of the GTR and Einstein's theory generally lacks any direct connection between the variation of matter and energy and curvature wave emission, the GTR cannot, in principle, explain a matter energy loss via gravitational radiation. Therefore a study of the motion of double systems and the determination of the possible energy lost by these systems via gravitational radiation would be an essential experimental test for the field theory of gravitation vis a vis the GTR. An experimental observation of the matter energy lost via gravitational radiation would unambiguously undermine the GTR and serve as corroboration of the ideas of the field theory of gravitation.

As has been shown in Sec. 1.17, the behavior pattern of the Universe during the early stages of its evolution according to the field theory of gravitation differs qualitatively from the respective description in the GTR. Since the earlier stages of the Universe evolution substantially influence the flux densities and spectral characteristics of the remnants of electromagnetic, neutrino, and gravitational radiation, then the predictions of the GTR and of the field theory of gravitation could be compared by measuring those characteristics.

Another effect due to strong gravitational fields is the description of the internal structure of superdense objects. The difference in describing the internal structure of stars in the field theory of grav-

itation and in the GTR should thus result in the different values for the limiting masses of stable stars.

There are also substantial differences between the field theory of gravitation and Einstein's GTR for the properties of gravitational waves when an external gravitational field is present.

Gravitational waves are generally called waves of the metric in the GTR and their theoretical investigation is usually carried out using the energy-momentum pseudotensors.

Yet for us this approach lacks any physical sense. The GTR's energy-momentum pseudotensors have, in principle, no relation to the existence of a gravitational field and so all the conclusions that are derived therefrom cannot reflect the essence of the problem. One may only speak about curvature waves in Einstein's theory, since the physical characteristic of the gravitational field in this theory is a curvature tensor. It is this tensor which enters the deviation equation (1.3) which forms the basis for the operation of any quadrupole mass detector of gravitational waves.

The presence of curvature waves in a particular domain of space-time is an explicit indication of the presence of gravitational waves in the region that are being emitted by some source. The curvature waves cannot be either produced or destroyed by transforming the coordinate frame. Yet metric waves cannot serve as a criterion of the presence of gravitational waves emitted by some source, since metric waves can be produced from a simple transformation of the coordinate frame.

There below we shall use GTR gravitational waves to imply curvature waves described by a tensor of the fourth rank, i.e. a curvature tensor.

In the GTR the inherent geometry for electromagnetic waves and for metric waves is a unified Riemannian geometry. Since curvature waves are expressed through the second derivatives of the transverse part of the metric waves, the Riemannian geometry is the inherent one for curvature waves as well. Hence the propagation of electromagnetic and curvature waves proceed in Einstein's theory in the same way, that is curvature waves, just like electromagnetic ones, are prone to a gravitational frequency shift $\delta\nu/\nu = U_1 - U_2$, they experience the same ray deflection $\delta\varphi = 4M/b$ and have equal propagation velocities resulting in the same delay times in external gravitational fields.

In the field theory of gravitation the inherent geometry for a gravitational field is the pseudo-Euclidean one, whereas matter is described by the effective Riemannian geometry. Hence in the field theory of gravitation external gravitational fields exert an influence only on electromagnetic waves. Gravitational waves propagate, in the field theory of gravitation, along the geodesic of pseudo-Euclidean space-time and experience no frequency red shift due to gravity, no deflection of their rays nor signal time delays in external gravitational

fields. The propagation velocity of gravitational waves in the field theory is independent of external gravitational fields.

These differences between the field theory of gravitation and Einstein's GTR in the properties of gravitational waves enable a number of experiments using weak gravitational waves to be proposed for which the two theories yield different predictions. The following realizations of such experiments are two main ones that are possible.

a. Experiments Using Laboratory Detectors of Gravitational Waves

It is anticipated [112-119] that in the near future it will be possible to detect, under laboratory conditions, gravitational radiation emitted by extraterrestrial sources. By using two or more gravitational wave detectors the angle by which a gravitational ray is bent in the weak gravitational field of the Sun could be measured [32]. The success of this experiment depend on whether the extraterrestrial source of gravitational waves is a source of electromagnetic radiation as well.

If the gravitational wave source does not emit electromagnetic waves, the experiment will be similar to measuring the bending of light ray. In this case, comparing the measured value of the gravitational ray's bending angle with suitable values predicted by the field theory of gravitation ($\delta\varphi = 0$) and by Einstein's GTR ($\delta\varphi = 4M/b$) will show how well the predictions of these theories agree with the experimental results.

If the gravitational wave source emits electromagnetic waves as well, the experimental scheme will be considerably simpler, since the "electromagnetic" and "gravitational" source images will always coincide in the GTR.

As to the field theory of gravitation, the picture will be slightly different. When the line connecting the gravitational wave source and an observer approaches the edge of the solar disc, the "gravitational" and "electromagnetic" source images will start to bifurcate, the "electromagnetic" image being observed further from the center of the Sun than the "gravitational" one. When the line that connects the observer and the gravitational wave source touches the solar disc's edge, the angular separation between the "electromagnetic" and "gravitational" source images will be maximum and equal to $\delta\varphi = 4M/b$. As soon as the solar disc overlaps the source-observer line, the "electromagnetic" image will vanish. Following the passage of this line across the solar disc the "electromagnetic" image will reappear. It will again be observed further away from the center of the Sun, by an angular separation $\delta\varphi = 4M/b$, than the "gravitational" image position appears to be. As the solar disc moves away from the source-observer line, the angular separation between the "electromagnetic" and "gravitational" images will reduce and as $b \to \infty$, both images will coincide.

b. Experiments Using Laboratory and Outer Space Detectors of Gravitational Waves

However, one would expect that rather powerful sources of weak gravitational waves are exceedingly rarely encountered under astronomical conditions. Hence it would appear that those sources of gravitational waves which could be registered on the Earth would be fairly remote from the Earth's orbital plane and, as a consequence, would not be overlapped by the solar disc. In this situation, it becomes essential to set up an experiment for which the field theory of gravitation and the GTR yield distinguishable predictions and which can use cosmic detectors instead of the laboratory detectors of gravitational waves.

Nowadays the problem with detecting gravitational waves is usually associated with the ability to detect them under laboratory conditions. However, we think that gravitational waves could be detectable via peculiar features of electromagnetic waves as well which emerge from an interaction of gravitational waves with the electromagnetic fields of astronomical objects, the field of a rotating neutron star, for example.

This new technique of detecting gravitational waves may be more effective in certain situations than the conventional laboratory detecting techniques, since neutron stars possess electromagnetic fields having strength unaccessible in the laboratory ($\mathscr{H} \sim 10^{12}$-10^{15} Oe) and fields of this strength extend to considerable distances ($r \sim 10^6$-10^8 cm), thus forming cosmic detector of gravitational waves.

Another advantage of the proposed new technique is that the electromagnetic waves which would emerge from the interaction could be registered using a modern radiotelescope that achieved a collecting area of 10^4 m^2. The development of laboratory detectors of gravitational waves with such a cross section is highly improbable even in the near future.

Depending on where the gravitational wave source and the cosmic detector are positioned relative to each other, the two following schemes are conceivable.

The first can be realized when the source of weak gravitational waves is located in the interior of a rotating neutron star.

Many neutron stars [120] have an electromagnetic field matching that of a rotating magnetic dipole, the magnetic axis forming an angle ψ_0 with respect to the axis of rotation. In addition, according to [121, 122], the photoproduction of gravitons in the core of the star in the Coulomb and magnetic fields of the particles which constitute the stellar mass as well as the magnetic field of the star as a whole takes place.

The stars are thus the sources of weak gravitational waves, gravitational radiation with potentially any frequency range depending on the initial photon spectrum.

Calculations show that in both the field theory of gravitation and Einstein's GTR the interaction of weak gravitational waves with the electromagnetic field of a rotating neutron star results in an electromagnetic wave which has a number of unique features (amplitude modulation, nonstandard polarization, subimpulse drift, etc.). Thus, an observer on Earth would deduce from these features and with a high confidence level that a given electromagnetic wave emerged from an interaction between a gravitational wave and the electromagnetic field of the rotating star. The amplitude of the resultant electromagnetic wave would differ slightly in these theories due to the differing influence of the static gravitational field on the interaction process, though this is generally unessential. Of greater importance for us is that the field theory of gravitation and Einstein's GTR predict that the electromagnetic and initial gravitational waves produced would subsequently propagate in the outer static gravitational field in different fashions.

According to Einstein's GTR, both waves would propagate along the same trajectories (rays) experiencing the same frequency red shifts due to gravity, encountering the same rays bending, and having the same group velocities. Hence an observer on the Earth, detecting these waves, must find that their frequencies are coinciding, they have the same pulse shape within a "window", there is no time delay between the arrival of the electromagnetic and gravitational pulses, and that the "electromagnetic" and "gravitational" source images coincide.

In the field theory of gravitation the electromagnetic wave produced, while propagating in the external gravitational field, will also be subjected to a frequency red shift due to gravity, have its rays bend, and find its group velocity depending on the external field potential. Yet in the field theory of gravitation gravitational waves are not influenced by gravitational field, hence they will propagate at constant velocity without changing the frequency or bending in the external gravitational field.

In this situation the terrestrial observer, detecting both waves, must discover that the electromagnetic spectrum is more shifted to the red relative to the gravitational one, that there is a time delay between the arrival of the gravitational pulses and the electromagnetic pulses within the "window". In addition, the "gravitational" and "electromagnetic" source images should not coincide in general.

Note that the results of this experiment would also yield some other important astrophysical data. Indeed, as has been shown in [123, 124], the rotation frequency ω_0 of the neutron star, the angle ψ_0 between the rotation axis and the magnetic moment of the star, and the angle θ between the rotation axis and the director vector towards the Earth could be defined from the frequency and depth of the amplitude modulation and from the polarization of the electromagnetic wave. By measuring the flux densities of the gravitational and

electromagnetic waves on the Earth, we could determine the conversion factor α and from that the value of the product of the magnetic field strength on the star's surface and its radius.

In addition, this experiment in the field theory of gravitation enables the difference of gravitational potentials between the observation point and the star's surface to be derived from the red shift of the electromagnetic spectrum relative to the gravitational one:

$$U_2 - U_1 = -\delta\nu/\nu.$$

By measuring the delay time ΔT between the arrival of the gravitational and electromagnetic pulses, one may determine the mean gravitational potential $\overline{U}$ along the propagation path of the electromagnetic wave:

$$\overline{U} = \frac{1}{L} \int U\, dL \simeq \Delta T/2L,$$

where L is the distance separating the neutron star and the Earth.

If the gravitational wave source is located outside the neutron star, the analysis of the observational data will be somewhat more sophisticated, since the source, the cosmic detector, and the laboratory detector will be located at the apexes of a triangle. However, some conclusions can be made in this case too about the properties of gravitational waves as a result of the observations and a set of data on the astronomical objects can be obtained.

Thus, in the near future, once laboratory detectors of gravitational waves have been developed, there will be the real possibility of checking the predictions of the field theory of gravitation against those of Einstein's GTR concerning the properties of gravitational waves in external gravitational fields.

1.19 CONCLUSION

We have considered the formulation of the field theory of gravitation that uses a symmetric tensor field of the second rank in a flat space-time. In this way the conventional concepts of the transfer of energy by physical fields have a rigorous meaning in the theory and a gravitational field, like all other physical fields, transfers a positive energy-momentum. The equations of motion of the matter are formulated in terms of an effective Riemannian space-time that has the metric tensor g_{ni}. This ensures that the inertial and gravitational masses of a point body are equal in the theory. Unifying the concepts of the gravitational field, to make it one that transfers energy and is similar to other physical fields, with the identity principle results in new gravitational field equations and alters our concept of space-time. The gravitational field equations within matter are nonlinear due to a nonlinear source dependence on the gravitational field's components. The matter appears in the gravitational

field equations as a source, the gravitational field itself being the source only so far as the expression $T^{ki}\partial g_{ki}/\partial f_{nm}$ depends on the gravitational field's components. The field equations are linear outside matter and so, due to gauge invariance, the gravitational field equations, which involve fourth-order partial derivatives, become second-order equations outside the matter.

A post-Newtonian approximation of the field theory of gravitation and an analysis of modern gravitational, experimental data demonstrate that the minimum-coupling field theory of gravitation enables all the available experimental facts to be described. There is no preferred rest system in the field theory of gravitation, since the pseudo-Euclidean space-time geometry is not a priori but rather the one inherent for all physical fields including the gravitational. Yet Riemannian space-time is an effective space-time for the motion of matter and it reflects only the action of a gravitational field on matter in pseudo-Euclidean space-time. Therefore, the field theory of gravitation does not fall in a class of the so-called bimetric theories of gravitation either in essence or in the form of its field equations.

In the field theory of gravitation, the energy-momentum tensor notion is common for all physical fields, hence the existence of curvature waves in the Riemannian space-time accounts for the energy-momentum transfer by gravitational waves in the pseudo-Euclidean space-time. As a consequence, the field theory of gravitation allows different energy-based computations to be carried out. In the field theory of gravitation, the radiative losses of weak waves by a slowly moving source are specified by the expression

$$-\frac{dE}{dt}=\frac{G}{45c^5}\dddot{D}_{\alpha\beta}^{2}. \tag{1.288}$$

As opposed to the field theory of gravitation, conservation laws in their common sense are absent in the GTR. This results in Einstein's theory only in integrals of motion which vanish. It appears to be impossible to calculate a source energy loss as well as determine the gravitational-wave energy fluxes in the GTR, since there are no GTR conservation laws which relate changes in the energy-momentum tensor of matter to the existence of curvature waves. Therefore formula (1.288) does not, in principle, follow from the GTR.

Equations of the field theory of gravitation differ from those of Einstein's theory in that they yield considerably different descriptions of the effects of strong gravitational fields as well as different properties of gravitational waves. These differences include the fact that according to the field theory, gravitational rays do not bend when passing near massive bodies which means that massive bodies do not focus the gravitational waves. In addition, the field theory, as opposed to Einstein's GTR, predicts that the frequency of a free gravitational wave emitted by some source only changes when the

source moves relative to the observer (the Doppler effect), since a red shift of free gravitational waves due to gravity is absent in vacuo.

In the field theory of gravitation as the Einstein's theory, the gravitational field of a non-static, spherically symmetric source is a static field outside matter. A non-stationary homogeneous Universe model in the field theory of gravitation describes the cosmological red shift and allows both monotonic and non-monotonic behavior. As opposed to Einstein's theory, the deceleration parameter is specified not only by the average density of the matter in the Universe, but also by the post-Newtonian coefficients. Hence in the field theory of gravitation there is none of the difficulties Einstein's GTR is encountering related to an insufficient average mass density needed to account for the observed deceleration parameter.

In the minimum-coupling, field theory of gravitation, gravitational attractive forces become repulsive when the gravitational potential is increased. Therefore, instead of the gravitational collapse of astronomical objects typical of the GTR, a new mechanism of energy release is described. Small radial perturbations of an astronomical object occurring at the critical value of the average gravitational potential will inevitably lead to an expansion of the matter accompanied by the ejection of part of the object's mass and an energy release.

References

1. A. A. Logunov and V. N. Folomeshkin, *Theor. and Math. Phys.*, **32**, No. 2, 653 (1977).
2. A. A. Logunov and V. N. Folomeshkin, *Theor. and Math. Phys.*, **32**, No. 2, 667 (1977).
3. A. A. Logunov and V. N. Folomeshkin, *Theor. and Math. Phys.*, **32**, No. 3, 749 (1977).
4. A. A. Logunov and V. N. Folomeshkin, *Theor. and Math. Phys.*, **33**, No. 2, 952 (1977).
5. V. I. Denisov and A. A. Logunov, *Theor. and Math. Phys.*, **43**, No. 2, (1980).
6. V. I. Denisov and A. A. Logunov, *Theor. and Math. Phys.*, **45**, No. 3, (1980).
7. A. A. Logunov and V. I. Denisov, in: *Abstr. Contrib. Papers 9th Intern. Conf. Gen. Relat.*, v. 3, Jena (1980), p. 184.
8. V. I. Denisov and A. A. Logunov, *Preprint IYaI AN SSSR* P-0206 (1981) (in Russian).
9. A. A. Vlasov and V. I. Denisov, *Preprint IFVE* OTF 81-141 (1981) (in Russian).
10. J. L. Synge, *Relativity: The General Theory*, North-Holland Publ., Amsterdam (1960), p. VIII.
11. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, Addison-Wesley Publ., Reading, Mass. (1962).

12a. A. Einstein, *Sitzungsber. preuss. Akad. Wiss.*, **1**, 167 (1918).
12b. A. Einstein, *Phys. Z.*, **15**, 108 (1914).
12c. A. Einstein, *Sitzungsber. preuss. Akad. Wiss.*, **1**, 167 (1918).
12d. A. Einstein. *Sitzungsber. preuss. Akad. Wiss.*, **1**, 450 (1918).
12e. A. Einstein, *Sitzungsber. preuss. Akad. Wiss.*, **1**, 459 (1918).

12f. A. Einstein, *Sitzungsber. preuss. Akad. Wiss.*, **1**, 164 (1918).
12g. A. Einstein, *Ann. Phys.*, 49, 809 (1916); in: The Principle of Relativity, Dover Publ., New York (1952), p. 151.
13. V. A. Fock, *The Theory of Space, Time, and Gravitation*, Pergamon Press, London (1959).
14. C. Møller, *The Theory of Relativity*, Clarendon Press, Oxford (1972), 557 p.
15. K. P. Stanyukovich, *Gravitational Field and Elementary Particles*, Nauka, Moscow (1965) (in Russian).
16. R. C. Tolman, *Relativity, Thermodynamics, and Cosmology*, Clarendon Press, Oxford (1969).
17. S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*, Wiley, New York (1972), 657 p.
18. A. S. Eddington, *The Mathematical Theory of Relativity*, Cambridge Univ. Press, Cambridge (1923).
19. D. Brill, S. Deser, and L. Faddeev, *Phys. Lett.*, **26A**, 538 (1968).
20. I. Choquet-Bruhat and J. E. Marsden, *Math. Phys.*, **51**, 283 (1976).
21. P. Schoen and S.-T. Yau, *Math. Phys.*, **65**, 45 (1979).
22. N. Hu, *Proc. Roy. Irish Acad.*, **A51**, 87 (1947).
23. S. Smith and P. Havas, *Phys. Rev.*, **B138**, No. 2, 495 (1965).
24. L. Infeld, *Ann. Phys.*, **6**, No. 4, 341 (1959).
25. A. Trautman, *Bull. Acad. Polon. Sci.*, **6**, No. 10, 627 (1958).
26. I. V. Sandina, PhD thesis, Alma-Ata (1972) (in Russian).
27. N. M. Petrova and I. V. Sandina, *Sov. Phys. Dokl.*, **19**, No. 7, 443 (1975).
28. A. A. Logunov et al., *Theor. and Math. Phys.*, **40**, No. 3 (1979).
29. A. A. Vlasov, V. I. Denisov, A. A. Logunov, and M. A. Mestvirishvili, *Theor. and Math. Phys.*, **43**, No. 2 (1980).
30. V. I. Denisov, A. A. Logunov, and M. A. Mestvirishvili, *Fiz. Elem. Chastits i Atom. Yadra*, **12**, No. 1, 5 (1981).
31. V. I. Denisov and A. A. Logunov, in: *Abstr. Contrib. Papers 9th Intern. Conf. Gen. Relat.*, v. 2, Jena (1980), p. 468.
32. V. I. Denisov and A. A. Logunov, *Preprint IYaI AN SSSR* P-0199 (1981) (in Russian).
33. V. I. Denisov and A. A. Logunov, *Abstr. Contrib. Papers 5th Sov. Grav. Conf.* [ed. A. A. Sokolov], Moscow University Press, Moscow (1981) p. 110 (in Russian).
34. N. I. Lobachevskii, *Collected Papers*, v. 2, Moscow (1949), p. 159 (in Russian).
35. A. Z. Petrov, *New Methods in the General Theory of Relativity*, Nauka, Moscow (1966) (in Russian).
36. H. Poincaré, *The Monist*, **15**, ser. 1, 1 (1905), cited in: L. P. Williams, *Relativity Theory: Its Origins and Impact on Modern Thought*, Wiley, New York (1968), p. 98.
37. H. Minkowski, in: *The Principle of Relativity*, Dover Publ., New York (1952), p. 75.
38. A. A. Logunov, *Voprosy Filosofii*, No. 6, 21 (1981).
39. A. A. Logunov, *Methodological Instructions for the Lectures in the Series "Fundamentals of the Theory of Relativity"*, MFTI, Moscow (1980) (in Russian).
40. P. K. Rashevskii, *Riemannian Geometry and Tensor Analysis*, Nauka, Moscow (1967) (in Russian).
41. A. Einstein, *Ann. Phys.*, 38, 355 (1912).
42. A. Einstein, *Phyz. Zs.*, **14**, 1249 (1913).
43. G. Nordström, *Ann. Phys.*, **40**, 856 (1913).
44. G. Nordström, *Ann. Phys.*, **42**, 533 (1913).
45. A. Einstein and A. D. Fokker, *Ann. Phys.*, **44**, 321 (1914).
46. A. N. Whitehead, *The Principle of Relativity*, Cambridge Press, Cambridge (1922).
47. D. E. Littlewood, *Proc. Cambridge Phil. Soc.*, 49, No. 1, 90 (1953).
48. A. Papapetrou, *Zs. Phys.*, 139, 518 (1954).

49. A. Papapetrou, *Math. Nachr.*, **12**, 129 (1954).
50. O. Bergman, *Amer. J. Phys.*, **24**, No. 1, 38 (1956).
51. R. H. Dicke, *Rev. Mod. Phys.*, **29**, No. 3, 363 (1957).
52. H. Yilmaz, *Phys. Rev.*, **111**, No. 5, 1417 (1958).
53. Ya. I. Pugachov, *Izv. VUZov*, Fizika, No. 6, 152 (1959).
54. W. Thirring, *Forschr. Phys.*, **7**, 79 (1959).
55. G. J. Whitrow and G. E. Morduch, *Nature*, **188**, No. 4753, 790 (1960).
56. H. Yilmaz, in: *Evidence for Gravitational Theories*, New York (1962), p. 34
57. H. Dehnen, *Ann. Phys.*, **13**, No. 3, 101 (1964).
58. R. Godelupi, *Note, recens. notiz.*, **19**, No. 2, 258 (1970).
59. M. Muraskin, *Ann. Phys.*, **59**, No. 1, 27 (1970).
60. R. V. Wagoner, *Phys. Rev.* **D 1**, No. 12, 3209 (1970).
61. A. Z. Petrov, *Sov. Phys. Dokl.*, **15**, No. 1, 40 (1970).
62. L. Brillouin, *Relativity Reexamined*, Acad. Press, New York (1971), 111 p.
63. W. T. Ni, *Astrophys. J.*, **176**, No. 3, 769 (1972).
64. N. Rosen, *Phys. Rev.* **D 3**, No. 10, 2317 (1971).
65. N. Rosen, *Gen. Relat. and Gravit.*, **4**, No. 6, 435 (1973).
66. N. Rosen, *Ann. Phys.*, **84**, No. 1, 455 (1974).
67. W. T. Ni, *Phys. Rev.* **D 7**, No. 10, 2880 (1973).
68. A. P. Lightman and D. L. Lee, *Phys. Rev.* **D 8**, No. 10, 3293 (1973).
69. H.-J. Treder, *Gravitationstheorie und Äquivalenzprinzip*, Berlin (1971).
70. W. Petry, *Ann. Inst. H. Poincaré*, **A22**, No. 4, 277 (1975).
71. W. Petry, *Ann. Phys.*, **33**, No. 1, 6 (1976).
72. Ch. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*, W. H. Freeman and Co., San Fransisco (1973), part VIII.
73. W. Thirring, *Ann. Phys.*, **16**, No. 1, 96 (1961).
74. Ya. B. Zel'dovich and I. D. Novikov, *The Theory of Gravitation and Evolution of Stars*, Nauka, Moscow (1971) (in Russian).
75. Ch. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*, W. H. Freeman and Co., San Fransisco (1973), part II.
76. N. N. Bogolyubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields*, Intersci. Publ., New York-London (1959).
77. V. I. Ogievetsky and I. V. Polubarinov, *Ann. Phys.*, **35**, 167 (1965).
78. C. Fronsdal, *Nuovo Cim.*, **9**, No. 2, 416 (1958).
79. K. J. Barnes, *J. Math. Phys.*, **6**, No. 5, 788 (1965).
80. V. S. Vladimirov, *Equations in Mathematical Physics*, Nauka, Moscow (1971) (in Russian).
81. C. M. Will, *Astrophys. J.*, **185**, No. 1, 31 (1973).
82. R. A. Hulse and J. H. Taylor, *Astrophys. J.*, **195**, No. 2, L51 (1975).
83. J. H. Taylor *et al.*, *Nature*, **277**, 437 (1979).
84. C. M. Will and K. Nordtvedt, *Astrophys. J.*, **177**, No. 3, 757 (1972).
85. C. M. Will, *Astrophys. J.*, **169**, No. 1, 125 (1971).
86. D. L. Lee *et al.*, *Phys. Rev.* **D 10**, No. 6, 1685 (1974).
87. O. S. Ivanitskaya, *The Lorentz Basis and Gravitational Effects in Einstein's Theory of Gravity*, Nauka i Tekhnika, Minsk (1979) (in Russian).
88. I. Shapiro, *Science*, **157**, No. 3790, 806 (1967).
89. I. Shapiro, *Phys. Rev. Lett.*, **13**, No. 26, 789 (1964).
90. R. D. Reasenberg *et al.*, *Astrophys. J. Lett.*, **234**, No. 3, 219 (1979).
91. R. F. Oconnel, *Gen. Relat. and Gravit.*, **3**, No. 2, 123 (1972).
92. N. P. Konoplyova, *Sov. Phys. Usp.*, **20**, No. 12, 973 (1977).
93. V. B. Braginskii and A. G. Polnaryov, *Pis'ma ZhETF*, **31**, No. 8, 444 (1980).
94. I. I. Gol'denblat and S. V. Ul'yanov, *Introduction to the Theory of Relativity and Its Applications in New Technology*, Nauka, Moscow (1979) (in Russian).
95. C. M. Will, in: *Proc. IX Texas Symp. Relat. Astrophys.*, Munich (1978), p. 1.
96. R. H. Dicke and H. M. Goldberg, *Phys. Rev. Lett.*, **18**, No. 9, 313 (1967).

97. I. Shapiro, *Gen. Relat. and Gravit.*, **3**, No. 2, 135 (1972).
98. K. Nordtvedt, *Phys. Rev.*, **170**, No. 5, 1186 (1968).
99. K. S. Thorne and C. M. Will, *Comments Astrophys. Space Phys.*, **2**, No. 1, 35 (1970).
100. K. Nordtvedt, *Phys. Rev. D.*, **7**, No. 8, 2347 (1973).
101. J. G. Williams *et al.*, *Phys. Rev. Lett.*, **36**, No. 11, 551 (1976).
102. I. Shapiro *et al.*, *Phys. Rev. Lett.*, 36, No. 11, 555 (1976).
103. C. M. Will, *Astrophys. J.*, **169**, No. 1, 141 (1971).
104. V. B. Braginskii and V. I. Panov, *ZhETF*, **61**. No. 3, 873 (1971).
105. V. I. Denisov and A. A. Logunov, *Theor. and Math. Phys.*, **48**, No. 3, (1981).
106. Ch. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*, W. H. Freeman and Co., San Fransisco (1973), p. 630-634.
107. J. Oort, *Distribution of Galaxies and Density in the Universe*, Brussels (1959), 183 p.
108. D. Ya. Martynov, *A General Astrophysics Course*, Nauka, Moscow (1971) (in Russian).
109. E. P. Hubble, *Proc. Nat. Acad. Sci.*, **15**, 169 (1929).
110. A. Friedman, *Phys. Z.*, **10**, No. 11, 377 (1922).
111. E. L. Turner, *Astrophys. J.*, **230**, No. 2, 291 (1979).
112. V. B. Braginskii and A. B. Manukin, *Measurements of Weak Forces in Physics Experiments*, Univ. Chicago Press, 1977.
113. N. B. Mitskevich, *Physical Fields in the General Theory of Relativity*, Nauka, Moscow (1969) (in Russian).
114. V. B. Braginskii and V. N. Rudenko, *Sov. Phys. Usp.*, **13**, 165 (1970).
115. V. D. Zakharov, *Gravitational Waves in Einstein's Theory of Gravity*, Nauka, Moscow (1972) (in Russian).
116. V. B. Braginskii and M. B. Menskii, *Sov. Phys. JETP Lett.*, **13**, No. 11, 417 (1971).
117. V. B. Braginskii *et al.*, *ZhETF*, **65**, No. 5, 1729 (1973).
118. I. M. Ternov and V. P. Khalilov, in: *Classical and Quantum Theory of Gravitation* (Abstr. Contrib. Papers All-Union Conf.), Minsk (1976), p. 135 (in Russian).
119. I. M. Ternov, V. P. Khalilov, and G. A. Chizhov, *ZhETF*, **68**, 6, 1969 (1975).
120. F. J. Dyson, *Neutron Stars and Pulsars*, Fermi Lectures 1970, Acad. Naz. Linc., Roma, 1971.
   D. Ter-Haar, *Pulsars*, Physics Reports, **3C**, No. 2, March 1972.
121. F. D. V. Gal'tsov, *ZhETF*, **67**, No. 2, 425 (1974)
122. G. Papini and S. R. Valluri, *Can. J. Phys.*, **53**, No. 20, 2306 (1975).
   G. Papini and S. R. Valluri, *Can. J. Phys.*, 53, No. 20, 2312 (1975).
123. V. I. Denisov, *ZhETF*, **74**, 401 (1978) (in Russian).
124. V. I. Denisov, *Izv. VUZov*, Fizika, No. 3, 152 (1978) (in Russian).
125. *Collected Papers of Academician A. N. Krylov*, v. 1 (ed. S. Ya. Shtraikh), AN SSSR Publ., Moscow (1951) (in Russian).
126. L. D. Faddeev, *Preprint LOMI AN SSSR* P-18-81 (1981) (in Russian).
127. L. P. Grishuk and Ya. B. Zel'dovich, *Preprint IKI AN SSSR* P-674 (1981) (in Russian).
128. A. Einstein, L. Infeld, and B. Hoffman, *Ann. Math.*, **39**, 65 (1938).
129. A. Einstein and L. Infeld, *Ann. Math.*, **41**, 455 (1940).
130. A. Einstein and L. Infeld, *Can. J. Math.*, **1**, 209 (1949).

2 Inclusive Processes and Strong Interaction Dynamics

A. A. Logunov, M. A. Mestvirishvili, and V. A. Petrov

2.1 INTRODUCTION

2.1.1 High Energy Physics

When people speak about high energy physics they mean the branch of physics that investigates the collision processes of the very fast particles, that are either produced in accelerators, or come from outer space.

The main goal of these investigations is to understand the fundamental principles of the surrounding world and to discover whether it is possible to reduce all its diversity to a few simple elements.

The problem is far from new, it dates as far back as philosophers of antiquity. Yet it was only early in the 20th century that atomic theory of matter structure became universally accepted. At this time the electron had been discovered and questions began to be asked about the structure of atoms *per se*. The first person to take the topic seriously in hand was E. Rutherford and in 1911 he and his colleagues succeeded when they discovered an atom's structure. Their conclusion came as a result of bombarding gold foil by α-particles during their experiments and at that time the energy of the α-particles they used ($\sim$10 MeV) seemed to be very high.

Rutherford's studies can in this way be taken as the beginning of high-energy physics bearing in mind that the term "high energy" itself does not imply the sort of extremely high energies that can be achieved today. Indeed, to advance further into the depths of the structure of matter required the development of accelerators and a whole store of other experimental techniques so that nowadays the particle energies attainable surpass the particle energies of Rutherford's experiments by many orders of magnitude.

2.1.2 Multiple Production of Particles

A distinctive feature of the physics of microworld is the quantum transmutation of elementary particles. Studying these phenomena is one of the main ways of getting to know the new forms of matter and conservation laws associated with them.

As a result of an interaction, the particles may merely be deflected off their initial trajectories, i.e. undergo an elastic scattering or the

interaction could result in the production of new particles as well, which are generally different from the initial ones, i.e. in an inelastic process. It is those inelastic processes for which the number of secondaries exceeds the number of primaries which are called *multiple production processes*. Other inelastic processes (where the number of particles does not change) involve charge exchanges.

The multiple production of particles is most prominent in processes in which the strong interaction predominates. In fact, the forces responsible for strong interaction (nuclear forces) have an extremely short range, of the order of 10^{-13} cm, and for nuclear interaction to occur the particles (i.e., protons) have to be brought closer together so that the distance between them is of the order of the nuclear forces range. Clearly, if the particle energy is high, then there may be instant when this energy may be concentrated within a very small volume. When the energy density inside the small volume becomes large, the quantum-mechanical laws dictate that the initial energy is spent not only to produce the same primaries (elastic scattering) but also to be transformed with a high probability into a large (when the energy is sufficiently high) number of secondaries.

The earliest theoretical studies of these characteristics of multiple production processes as the angle of emission and energy distributions of secondaries and an average number of secondaries date back to the late forties and early fifties. These were initiated by discoveries, during cosmic ray experiments of events which lead to an enormous number of secondaries.

W. Heisenberg's work must be quoted among these, since he suggested that the momentum distribution of secondaries at high energy is governed by the equations of classical nonlinear field theory and obtained asymptotic expressions for these distributions. In 1949 E. Fermi advanced a statistical model of multiple particle production assuming that in a system of strongly interacting particles a thermodynamic equilibrium is rapidly reached as soon as a collision has taken place. It is after this that secondaries are produced according to the laws of black body radiation.

Pomeranchuk and Landau later advanced their own versions of the statistical approach to the problem of multiple production of particles. A prediction is peculiar to all these (and subsequent) statistical type theories, i.e. the mean multiplicity of secondaries will grow quickly (power-like) as the initial energy increases.

Later studies of multiple production processes used other approaches as well, e.g., the multiperipheral model. In contrast to the statistical theories, the mean multiplicity raises rather slowly (logarithmically).

2.1.3 Inclusive Processes

At comparatively low energies the fraction of multiple production processes is insignificant but rises noticeably with energy. Even this fact shows how very important multiple processes are when studying strong interactions. It is therefore by no accident that the current research on multiple processes comprises an entire trend in the physics of high energy particles.

Experimental and theoretical studies of the multiple production of particles are very difficult because of an abrupt jump in the number of degrees of freedom and because the complexity of the kinematics increases as the number of secondaries grows. In fact, while the amplitude of a $2 \to 2$ process depends on two independent variables, the amplitude of a $2 \to n$ process would depend on $3n - 4$ independent variables, i.e. the number of independent variables rises rather rapidly with the number of particles.

The most thoroughly investigated processes, i.e. elastic scattering and charge exchange (the binary processes) leave the number of particles unchanged. In binary processes it is sufficient to follow the momentum of only one of the final particles since the momentum of the second particle is automatically known due to the energy-momentum conservation laws.

How full a kinematic description of amplitudes is necessary for the basic features of multiple production processes to be understood? Could any process be studied in the same way a binary process is, i.e. by tracing only one of the final particles without being concerned with others? These questions were answered in 1967. It has been shown [1, 2] that essential information about the interaction dynamics of particles can be obtained, once the characteristics of only one or two specified particles are studied in the final state so long as all the processes giving rise to these given particles are included. The same works introduced total differential cross sections to describe the production of the given particles in all the possible processes which contain these particles in the final state. The first high energy upper bounds were also obtained for such cross sections within the framework of axiomatic quantum field theory. All these processes that incorporate an arbitrary number of other particles allowable by conservation laws in addition to the given particles became known as inclusive processes and the relevant cross sections, as inclusive cross sections following a proposal by Feynman (1969).

The concept of inclusive process enables what at first sight may seem absolutely different processes, for example multiple production in hadron collisions and deep inelastic lepton-proton scatterings, to be considered from a single standpoint. The inclusive approach proved to be well suited for describing multiple production processes irrespective of the type of the colliding particles. It aided the development of the physics of elementary particles at high energies in many ways.

When the inclusive processes in experiments performed in 1968 at the Serpukhov 76 GeV accelerator were studied, a law of cross section scaling invariance was discovered, which was absolutely new for the microworld. It meant that the relative high energy production cross sections of strongly interacting particles (the hadrons) were universal functions of the reduced momentum $p/p_{\max}$, i.e. independent of the energy of primaries [3].

Following the experimental discovery of scaling a vast number of papers appeared which described this fascinating phenomenon more or less successfully. First of all, these were the papers by Yang *et al.* and Feynman [4] who advanced the hypotheses of limiting fragmentation (Yang *et al.*) and scale invariance (Feynman) based on a generalization of the experimental observations. The discovery of scaling gave rise to the advent of the parton model (Feynman) which can visualize a fast hadron as a set of weakly interacting composite parts, i.e. the partons. Now the parton model has gained wide recognition as a theory of quarks bound by the Yang-Mills fields and enables the basic regularities of deep inelastic processes to be understood quite successfully. Essentially, it was the study of inclusive lepton-hadron processes which stimulated the investigations of non-Abelian gauge theories and brought about the discovery of one interesting phenomenon, i.e. asymptotic freedom. Certain progress was made by Mueller [5] in the study of pure hadronic inclusive processes when he applied the complex angular momenta method to the inclusive processes.

By now the general picture of the dynamics governing the behavior of inclusive cross sections looks like the following.

The "hard" processes (deep inelastic scattering, electron-positron annihilation, production of massive lepton pairs, and processes of inclusive production of hadrons with high transverse momenta) are rather well understood and the energy evolution of inclusive cross sections is quantitatively described. However, the study of those details of multiple production that are related to the transitions of partons (quarks and gluons) to hadron states has met a number of significant difficulties. The same is characteristic of the "soft" processes, e.g. those for which the detected particles possess a low transverse momentum. Here one has to resort to phenomenological constructions which do not follow directly from the fundamentals of quantum field theory, for example, quantum chromodynamics (QCD).

Under these conditions methods based on the fundamental principles of quantum field theory rate highly in importance. As mentioned above, the pioneering results along these lines were obtained in [1, 2]. A set of new corollaries was later produced in the form of various bounds of the behavior of high energy inclusive distributions which follow from microcausality, spectrality, and unitarity [6]. Of particular importance is a conclusion on the universal char-

acter of structure functions [7] which describe processes of differing origin. It can be treated as a common theoretical basis for the models of the parton type.

It is quite evident that by limiting ourselves to inclusive processes we pay for the description's simplicity by losing a major portion of the information contained in exclusive processes[1]. However, what information about the final states do we need? For an analysis using the quantum theory we would be interested in the mean values of dynamic variables. For example, the mean value of the secondary particle's energy as any other additive quantity is given by an inclusive cross section. Binary quantities, for example the emission angle between two secondaries, must be used with the inclusive cross sections which describe two particle production, etc.

The situation here is absolutely analogous to one in statistical physics. In due course Bogolyubov introduced the so-called statistical operators of particle complexes which describe a given system of particles that is much smaller than the complete system. It turned out that these quantities were sufficient for the mean values of the single particle and dynamic binary additive variables that govern the global thermodynamic characteristics of the system to be computed. It is hoped that the use in high energy physics of inclusive descriptions of multiparticle processes can yield sufficiently rich information about the way the multiple production of particles takes place as a whole, without needing specific details of each individual process.

2.1.4 The Main Discoveries Made Whilst Investigating Inclusive Processes

Experimental studies of inclusive processes rather soon revealed new specific phenomena. As stated above, the phenomenon of scale invariance was discovered in Serpukhov in 1968 and was later confirmed in other scientific centers.

About the same time in Stanford experiments on the deep inelastic scattering of electrons by protons were in progress at SLAC. The relevant cross sections of the deep inelastic processes also revealed a scaling behavior, since the structure functions which depend on energy transferred and on the square of momentum transferred from the electrons to the protons proved to be functions of a ratio of these quantities alone [8]. Later on other processes were investigated as well, viz. the deep inelastic scattering of neutrinos, antineutrinos, and muon by protons, the hadron production in electron-positron annihilation, and the production of massive lepton pairs in hadron-

---

[1] An exclusive process is one for which the characteristics of all the particles in the final state are given.

hadron collisions. In all these processes scaling regularities were discovered.

At last in 1972 at CERN a strong discrepancy was revealed between the experimental data and extrapolations from the low transverse momenta range [9] when the inclusive production processes of hadrons with high transverse momenta were studied. Within this range (the high transverse momenta), a scaling dependence on the transverse momentum was found to take place.

Thus, studying inclusive processes proved to be extremely fruitful and brought to light a number of scaling regularities that had not been observed in individual inclusive channels before. This, in turn, enabled further understanding of strong interaction dynamics and the composite structure of elementary particles.

For the first time the methods of axiomatic quantum field theory were successful in analyzing rigorously multiple production processes and in obtaining general limits on a feasible form of the asymptotic behavior of the physical characteristics of high energy inclusive processes.

An entire set of new trends and concepts in high energy physics was established within the framework of the inclusive approach and various models of multiple production processes were developed, especially those associated with the internal structure of particles. Note in this connection that inclusive processes can be very conveniently applied to the parton model since summing over the undetectable hadron systems can give results that in many cases are weakly dependent on the unknown mechanisms of quark-gluon confinement.

At the same time it should be borne in mind that research using the inclusive approach is still at the development stage, and there is a vast number of very interesting and important problems whose solution is required by present day high energy physics. For example, the relations between various inclusive processes must be studied in detail in order to establish a more clear-cut relationship between the physical characteristics of the inclusive processes observed and the internal parameters of the particles which describe the particles' structure.

2.2 FUNDAMENTALS OF INCLUSIVE PROCESSES

When the particles a and b collide, various reactions allowed by conservation laws and selection rules can take place. Given an initial state, any combination of two or more particles in the final state will be called a collision channel. For example, the scattering process

$$a + b \to a + b$$

proceeding at any of the energies of the system (a, b) is due to the elastic channel. All remaining reactions will lead to inelastic channels.

A channel is considered "open" when at the prescribed energy of the system (a, b) the reaction corresponding to the given channel is allowed by the conservation laws and selection rules.

Each open channel can be given uniquely by an integer-valued "vector identifier" of the channel:

$$n = (n_1, n_2, \ldots, n_j).$$

This vector corresponds to the reaction

$$a + b \rightarrow n_1 c_1 + n_2 c_2 + \ldots + n_j c_j, \tag{2.1}$$

where c_i are sorts of the particles (having some arbitrary yet fixed ordering), and n_i are the numbers of the particles of sort c_i.

Reaction (2.1) may be written symbolically in the following form:

$$a + b \rightarrow nC,$$

where C is a "vector" whose components correspond to all possible particle sorts:

$$C = (c_1, c_2, \ldots).$$

It is clear that at the given energy the channel vector identifier "cuts off" only a finite number of elements from C. The channel identifier components are not independent, since they are connected by the linear relationships which express the conservation laws of quantum numbers.

By definition, $n_i \geqslant 0$, with zero n_i components signifying that the channel does not contain the particles of sort c_i. In addition, the n_i are limited by the condition

$$\sum_i m_i n_i \leqslant \sqrt{s} \equiv (E_a + E_b)_{c.m.s.}$$

(m_i being the mass of a particle of sort c_i) following from the energy-momentum conservation law: $k_a + k_b = k_{1c_1} + \ldots + k_{n_j c_j}$ and the inequality $E = \sqrt{k^2 + m^2} \geqslant m$.

Channel n is open if

$$s = s_n = (\sum_i m_i n_i)^2,$$

i.e. s_n is the threshold of a channel n.

To shorten records later on we will introduce the following notations:

$$|n| = n_1 + \ldots + n_j,$$
$$n! = n_1! \ \ldots \ n_j!,$$
$$\mathbf{K}_n = (\mathbf{k}_{1c_1}, \ldots, \mathbf{k}_{n_1c_1}; \ldots, \mathbf{k}_{1c_j}, \ \ldots, \mathbf{k}_{n_jc_j}),$$
$$d\tau^n = \prod_{i=1}^{j} \ \prod_{l=1}^{n_i} d^3k_{lc_i}; \quad d^3k_{lc_i} = d\mathbf{k}_{lc_i}/(2\pi)^3\, 2E_{lc_i}.$$

Here k_{lc_i} is 4-momentum of the lth particle of sort c_i. Given that some of the components of the vector n vanish, the relevant differentials will be absent in $d\tau^n$.

In the quantum theory physical processes are described by the amplitudes of the transitions that are elements of the S matrix by definition. The amplitude of a transition from the initial state α to the final state ω is generally expressed in the form

$$<\omega|S|\alpha> = <\omega|\alpha> + \text{[terms containing free paths of the system from } \alpha \text{ to } \omega] + i\,(2\pi)^4 \delta\,(P_\omega - P_\alpha)\, T_{\omega\alpha}, \quad (2.2)$$

where P_α (P_ω) is the total 4-momentum of the initial (final) state. The last term in (2.2) corresponds to the most interesting case, when all the particles take part in the interaction.

Later on we will study the processes having two particles in the final state more generally, i.e. with the bracketed term in (2.2) being absent.

Process (2.1) where all the characteristics of all participating particles are known is called exclusive.

The square of the modulus of the amplitude of channel n is[2]

$$T^n_{ab} = \langle \mathbf{K}_n | T | \mathbf{k}_a, \mathbf{k}_b \rangle \quad (2.3)$$

and according to quantum-mechanical rules, this describes a creation probability density of the channel n:

$$P^n_{ab}(\mathbf{K}_n; \mathbf{k}_a, \mathbf{k}_b) = \frac{(2\pi)^4 \delta \left(k_a + k_b - \sum_{i=1}^{j} \sum_{l=1}^{n_i} k_{lc_i}\right) |T^n_{ab}|^2}{\int d\tau^n (2\pi)^4 \delta \left(k_a + k_b - \sum_{i=1}^{j} \sum_{l=1}^{n_i} k_{lc_i}\right) |T^n_{ab}|^2}.$$

It is evident that

$$\int d\tau^n P^n_{ab}(\mathbf{K}_n; \mathbf{k}_a, \mathbf{k}_b) = 1.$$

---

[2] Relation (2.3) is the definition of the operator T.

The probability that the channel n comes into existence relative to all open channels is given by the expression

$$w_n = \frac{\int d\tau^n (2\pi)^4 \delta\left(k_a + k_b - \sum_i \sum_l \dot{k}_{lc_i}\right) |T_{ab}^n|^2}{\sum_m \int d\tau^m (2\pi)^4 \delta\left(k_a + k_b - \sum_i \sum_l k_{lc_i}\right) |T_{ab}^m|^2},$$

and the relevant probability density by

$$\frac{dw_n}{d\tau^n} = \frac{(2\pi)^4 \delta\left(k_a + k_b - \sum_{i,l} k_{lc_i}\right) |T_{ab}^n|^2}{\sum_m \int d\tau^m (2\pi)^4 \delta\left(k_a + k_b - \sum_{i,l} k_{lc_i}\right) |T_{ab}^m|^2}.$$

It is evident that

$$\sum_n w_n = 1,$$

and

$$\frac{1}{w_n} \frac{dw_n}{d\tau^n} = P_{ab}^n (\mathbf{K}_n;\ \mathbf{k}_a,\ \mathbf{k}_b).$$

The quantities P_{ab}^n, w_n, $dw_n/d\tau^n$ enable mean values of momenta, numbers of particl es, etc., to be determined.

For example, the mean value of the νth momentum component of a particle of sort c_i in the channel n is

$$\langle k_{c_i}^\nu \rangle_n = \int d\tau^n k_{c_i}^\nu P_{ab}^n .$$

The mean number of particles of sort c_i is evidently

$$\langle n_{c_i} \rangle = \sum_n n_i w_n.$$

Experimental data are usually given in terms of effective cross sections. We shall hereafter give expressions for the various cross sections as well as their relation to the above probabilities.

1. The *differential cross section of the exclusive process* (2.1) is

$$\frac{d\sigma_{ab}^n}{d\tau^n} = \frac{1}{\lambda^{1/2}(s,\ m_a^2,\ m_b^2)} (2\pi)^4 \delta\left(k_a + k_b - \sum_i \sum_l k_{lc_i}\right) |T_{ab}^n|^2,$$

$$\frac{1}{4} \lambda (x,\ y,\ z) = x^2 + y^2 + z^2 - 2xy - 2xz - 2yz.$$

2. The *partially integrated exclusive cross section* is

$$\frac{d\sigma_{ab}^n}{d\xi_1 \ldots d\xi_l} = \int d\tau^n \prod_{i=1}^{l} \delta(\xi_i - g_i(k)) \frac{d\sigma_{ab}^n}{d\tau^n},$$

where $g_i(k) = g_i(\mathbf{k}_a, \mathbf{k}_b; \mathbf{K}_n)$ are some functions of the momenta of the particles participating in process (2.1). For example,

$$\frac{d\sigma^n_{ab}}{dt_{ac}} = \int d\tau^n\, \delta(t_{ac} - (k_a - k_c)^2)\, \frac{d\sigma^n_{ab}}{d\tau^n}\,; \quad c \in C.$$

3. The *total cross section of the exclusive process* (2.1) is

$$\sigma^n_{ab} = \int d\tau^n \frac{d\sigma^n_{ab}}{d\tau^n} = \int d\sigma^n_{ab}\,,$$

Proceeding from $d\sigma^n_{ab}/d\tau^n$, one can define differential cross sections relative to a smaller number of variables, for example,

$$d\sigma^n_{ab}/d\tau^m = \int d\tau^{n-m}\, d\sigma^n_{ab}/d\tau^n, \tag{2.4}$$

where $m = (m_1, \ldots, m_j)$ is an integer vector such that $0 \leqslant m_i \leqslant n_i$. For vectors such as these we can write

$$m \leqslant n.$$

Had we measured exclusive cross sections for all the open channels at a given energy, we would have all the accessible information about the processes generated during the collisions between the particles a and b.

However, some of the global characteristics of the processes given birth to by the collision between the particles a and b do not need detailed analyses of each channel. Let us follow a fixed set of particles $mC = m_1c_1 + \ldots + m_jc_j$ in the final states of the processes that take place during the collisions of particles a and b. A combination of these processes is called an inclusive process and is written symbolically as

$$a + b \to mC + X, \tag{2.5}$$

where X denotes all the remaining particles (not belonging to mC) in all the channels which contribute to the creation of mC.

4. The *differential cross section of the inclusive process* (2.5) is defined in the following way:

$$\frac{d\sigma_{ab\to mCX}}{d\tau^m} = \sum_{n\geqslant m} \frac{d\sigma^n_{ab}}{d\tau^m}.$$

5. By integrating the differential cross section, one obtains the *total cross section of process* (2.5):

$$\sigma_{ab\to mCX} = \sum_{n\geqslant m} \sigma^n_{ab}.$$

6. The *total cross section of the process* $a + b \to X$, *when* $m = (0, 0, \ldots 0)$, is

$$\sigma_{ab\to X} \equiv \sigma^{\text{tot}}_{ab} = \sum_n \sigma^n_{ab}.$$

7. A *partially integrated cross section* is defined in a way similar to that for exclusive channels:

$$\frac{d\sigma_{ab\to mCX}}{d\xi_1 \ldots d\xi_l} = \int d\tau^m \prod_{i=1}^{l} \delta(\xi_i - g_i(k)) \frac{d\sigma_{ab\to mCX}}{d\tau^m},$$

$$g_i = g_i(\mathbf{k}_a, \mathbf{k}_b; \mathbf{K}_m).$$

The definitions of these cross sections allow their relationship with the probabilistic distributions given above to be established easily:

$$P^n_{ab} = \frac{1}{\sigma^n_{ab}} \frac{d\sigma^n_{ab}}{d\tau^n}.$$

8. The quantity

$$\rho_{ab\to mCX} = \frac{1}{\sigma_{ab\to mCX}} \frac{d\sigma_{ab\to mCX}}{d\tau^m} = \sum_{n\geqslant m} \frac{dw_n}{d\tau^m}$$

is obviously the *probability density of the momenta distribution for the particles of the system mC*, which is detected in the inclusive process (2.5), with

$$\int d\tau^m \rho_{ab\to mCX} = 1.$$

For example, $\rho_{ab\to cX}$, $c \in C$ defines the mean value of the νth component of a detected particle of sort c:

$$\langle k^\nu_c \rangle = \int d^3k_c k^\nu_c \rho_{ab\to cX}.$$

9. In a fixed channel one is able to define the *mean density of the particles of sort c in a momentum phase space* by averaging the true density of the particles of sort c:

$$\sum_{l=1}^{n_c} (2\pi)^3 2E_c(\mathbf{k}_{lc}) \delta(\mathbf{k}_c - \mathbf{k}_{lc})$$

with the weight P^n_{ab}:

$$\frac{dn_c}{d^3k_c} = \left\langle \sum_{l=1}^{n_c} (2\pi)^3 2E_{lc} \delta(\mathbf{k}_c - \mathbf{k}_{lc}) \right\rangle$$

$$= \int d\tau^n \sum_{l=1}^{n_c} (2\pi)^3 2E_{lc} \delta(\mathbf{k}_c - \mathbf{k}_{lc}) P^n_{ab}(\mathbf{K}_n; \mathbf{k}_a, \mathbf{k}_b)$$

$$= n_c \int d\tau^n (2\pi)^3 2E_c \delta(\mathbf{k}_c - \mathbf{k}_{1c}) P^n_{ab}(\mathbf{K}_n; \mathbf{k}_a, \mathbf{k}_b).$$

The last identity follows from the symmetry of P^n_{ab} relative to the momenta permutations of the particles of one sort.

It can be shown that

$$\int d^3k_c \frac{dn_c}{d^3k_c} = n_c.$$

The higher distributions are likewise defined as

$$\frac{d\,[n!/(n-m)!]}{d\tau^m} = \frac{n!}{(n-m)!} \int d\tau^{n-m} P^n_{ab},$$

$$\int d\tau^m \frac{d\,[n!/(n-m)!]}{d\tau^m} = \frac{n!}{(n-m)!} \equiv \frac{n_1!}{(n_1-m_1)!} \cdots \frac{n_j!}{(n_j-m_j)!}.$$

If the mean density of the number of the particles of sort c is required without reference to the channel they are created in, dn_c/d^3k_c has to be averaged with a given channel weight w_n. Then the mean density of the particles of sort c will read as

$$\frac{d\langle n_c\rangle}{d^3k_c} = \sum_n \frac{dn_c}{d^3k_c} w_n = \frac{1}{\sigma_{ab\to cX}} \sum_n n_c \frac{d\sigma^n_{ab}}{d^3k_c}.$$

Obviously

$$\int d^3k_c \frac{d\langle n_c\rangle}{d^3k_c} = \sum_n n_c w_n = \langle n_c\rangle.$$

Likewise,

$$\frac{d\langle [n!/(n-m)!]\rangle}{d\tau^m} = \sum_n \frac{d\,[n!/(n-m)!]}{d\tau^m} w_n,$$

$$\int d\tau^m \frac{d\langle [n!/(n-m)!]\rangle}{d\tau^m} = \left\langle \frac{n!}{(n-m)!} \right\rangle$$

$$\equiv \left\langle \frac{n_1!}{(n_1-m_1)!} \cdots \frac{n_j!}{(n_j-m_j)!} \right\rangle.$$

Defining the quantity $d\langle [n!/(n-m)!]\rangle/d\tau^m$ infers its relation to the differential cross sectinos:

$$\frac{d\langle [n!/(n-m)!]\rangle}{d\tau^m} = \frac{1}{\sigma_{ab\to mCX}} \sum_n \frac{n!}{(n-m)!} \frac{d\sigma^n_{ab}}{d\tau^m}.$$

10. The *inclusive spectrum of the process* is thus defined as

$$f_{ab\to mCX} = \sum_n \frac{n!}{(n-m)!} \frac{d\sigma^n_{ab}}{d\tau^m}.$$

with the normalization

$$\int d\tau^m f_{ab\to mCX} = \left\langle \frac{n!}{(n-m)!} \right\rangle \sigma_{ab\to mCX}.$$

11. An *inclusive*, *partially integrated spectrum* is defined in the same way as the partially integrated differential cross section:

$$\bar{f}_{ab\to mcX}(\xi_1, \ldots, \xi_l) = \int d\tau^m \prod_{i=1}^{l} \delta(\xi_i - g_i(k))\, f_{ab\to mcX}.$$

As regards the experimental studies of the differential cross sections and spectra of the inclusive processes, some explanations are appropriate.

Since particles of the same sort are indistinguishable, when the exclusive cross sections $d\sigma_{ab}^n/d\tau^n$ are measured, it is actually the quantity $n!\, d\sigma_{ab}^n/d\tau^n$ that is measured. Since T_{ab}^n has the form

$$T_{ab}^n = \langle \mathbf{K}_n \mid T \mid \mathbf{k}_a, \mathbf{k}_b \rangle$$
$$= \langle 0 \mid \prod_{i=1}^{j} \prod_{l=1}^{n_i} a_i(\mathbf{k}_{lc_i})\, T \mid \mathbf{k}_a, \mathbf{k}_b \rangle / \sqrt{n!},$$

then we can say that $n! d\sigma_{ab}^n/d\tau^n$ "measures" the non-normalized final states $\langle 0 \mid \prod_{i,l} a_i(\mathbf{k}_{lc_i})$. Here $a_i(\mathbf{k}_{lc_i})$ is the destruction operator of a particle of sort c_i having a momentum $\mathbf{k}_{lc_i}$.

If an integration is carried out over a part of the momenta, all the aforementioned will be related to the particles whose momenta remain unintegrated. In particular, measuring the distributions of particles of sort c in channel n yields, in fact the quantity $n_c d\sigma_{ab}^n/d^3k_c$ which refers to the state $\langle \mathbf{k}_{1c_1}, \ldots, \not{\mathbf{k}}_c, \ldots, \mathbf{k}_{n_jc_j} \mid a_c(\mathbf{k}_c)$.[3] Whence it is apparent that in order to measure an inclusive spectrum $f_{ab\to cX} = \sum_n n_c d\sigma_{ab}^n/d^3k_c$ the particular channel where the detected particle is created does not need to be known, i.e. the spectrum is measured in a straightforward fashion. Yet if the cross section $d\sigma_{ab\to cX}/d^2k_c = \sum_n d\sigma_{ab}^n/d^3k_c$ is to be defined, the production cross section of a particle must be initially measured in each channel for $d\sigma_{ab\to cX}/d^3k_c$ to be reconstructed the result of the measurement having been divided by n_c.

In this context inclusive spectra are simpler to measure than the relevant differential cross sections. This seems to be the explanation for the overwhelming predominance of experimental data just on inclusive spectra.

It should, however, be stressed that measuring differential inclusive cross sections is absolutely necessary if, for example, we want

---

[3] $\langle \mathbf{k}_{1c_1}, \ldots, \not{\mathbf{k}}_c, \ldots, \mathbf{k}_{n_jc_j} \mid$ denotes a state in which particle of sort c and momentum $\mathbf{k}_c$ are absent.

to know the mean momentum of a particle detected in the inclusive process $a + b \to c + X$. Indeed, by definition,

$$\langle k_c^\nu \rangle = \frac{1}{\sigma_{ab \to cX}} \int d^3k_c k_c^\nu \frac{d\sigma_{ab \to cX}}{d^3k_c}. \tag{2.6}$$

Using the inclusive spectrum, one can define the quantity

$$\frac{1}{\sigma_{ab \to cX}} \int d^3k_c k_c^\nu f_{ab \to cX} \tag{2.7}$$

to be the total mean momentum carried off by all the particles of sort c. Sometimes fallacious statements claiming that the mean particle momentum is quantity (2.7) divided by $\langle n_c \rangle$ can be found. In fact, we have

$$\frac{1}{\langle n_c \rangle} \frac{1}{\sigma_{ab \to cX}} \int d^3k_c k_c^\nu f_{ab \to cX} = \sum_n \langle k_c^\nu \rangle_n \frac{n_c}{\langle n_c \rangle} w_n.$$

Otherwise it follows from definition (2.6) that

$$\langle k_c^\nu \rangle = \sum_n \langle k_c^\nu \rangle_n w_n.$$

Obviously, these are absolutely different things provided that one does not require the condition $w_n \sim \delta_{n_c, \langle n_c \rangle}$ to be met.

In addition, the differential cross section of an inclusive process will later be shown to be more directly related to the unitarity condition, this being very important for establishing its analytical properties for angular variables.

The "global" characteristics, i.e. the differential cross sections and spectra of inclusive processes, are extremely convenient and simple both in their theoretical and experimental aspects, but they do result in a certain loss of information and give only an indication of how multiple production processes proceed "on the average".

Experimental data have demonstrated that the creation of a particle in various domains of the phase volume is dictated by the different mechanisms, whether they are the collective effects like pionization and diffractive dissociation, or, for example the elementary acts of hard collisions in inclusive processes with high transverse momenta. And the respective behavior of inclusive spectra and differential inclusive cross sections in terms of the detected particles' momenta and total energy is absolutely different for each of the various domains.

For example, it is customary to assume that in processes which produce a hadron with a high transverse momentum we are generally dealing with events that are taking place in the short-range collisions of primaries. The momenta distributions and mean number of particles created in such collisions are naturally expected to be appreciably different from the case of long-range collisions. Clearly

the "global" characteristics *per se* cannot yield sufficiently full information on what takes place under such specially selected conditions, since integrating over the momenta of undetected particles can completely "smooth away" the specific peculiarities of the events being singled out.

Thus, there is a necessity for new characteristics to be introduced that are free of the aforementioned disadvantages of the ordinary inclusive distributions.

12. *Associated (conditional) inclusive distributions.* In probability theory so-called conditional distributions are considered along with the ordinary probability distributions of random quantities. For example, let $F(x, y)$ be the probability distribution for two random quantities x and y, normalized to unity:

$$\int dx\, dy\, F(x, y) = 1.$$

Assume we are interested in a probability distribution for the quantity x given y changes within an interval Δ:

$$[y_0 - \Delta y, \quad y_0 + \Delta y].$$

Then a conditional distribution meets our requirements, viz.

$$G(x \mid \Delta) = \frac{\int_\Delta dy\, F(x, y)}{\int dx \int_\Delta dy\, F(x, y)},$$

where obviously

$$\int dx\, G(x \mid \Delta) = 1.$$

Reducing Δy to zero, we get the conditional distribution

$$G(x \mid y) = \frac{F(x, y)}{\int dx\, F(x, y)},$$

which is local with respect to the quantity y that "prescribes a condition".

Expanding Δ to cover the whole range of y leads us to the distribution

$$G(x) = \frac{\int dy\, F(x, y)}{\int dx\, dy\, F(x, y)} = \int dy\, F(x, y)$$

of the ordinary, "non-conditional" type.

Since inclusive distributions differ from the above example only by the concrete interpretation of the random quantities (momenta, multiplicities), we will take the example as a basis for determining the characteristics we need.

Consider the quantity

$$(2\pi)^3 2E_c \int\limits_{k_d \in V} d\tau^n \, \delta(\mathbf{k}_c - \mathbf{k}_{1c}) \, P^n_{ab}(\mathbf{K}_n;\ \mathbf{k}_a, \mathbf{k}_b); \quad c,\ d \in C,$$

where the integration with respect to the momentum $\mathbf{k}_d$ is performed only over a certain part V of an admissible phase volume. The quantity

$$P^n_{ab}(\mathbf{k}_c \mid V,\ d) = \frac{\int\limits_{k_d \in V} d\tau^n \, \delta(\mathbf{k}_c - \mathbf{k}_{1c}) (2\pi)^3 2E_c P^n_{ab}(\mathbf{K}_n;\ \mathbf{k}_a,\ \mathbf{k}_b)}{\int\limits_{k_d \in V} d\tau^n P^n_{ab}}$$

yields a conditional distribution over the momenta of one of the final particles c in the channel n. The condition is that at least one of the particles d, other than the detected one must not leave the phase volume domain. In terms of the cross sections we clearly have

$$P^n_{ab}(\mathbf{k}_c \mid V,\ d) = \frac{\int\limits_{(V)} d^3k_d \, d\sigma^n_{ab}/d^3k_c \, d^3k_d}{\sigma^n_{ab}(s \mid V,\ d)},$$

where

$$\sigma^n_{ab}(s \mid V,\ d;\ c) = \int\limits_{k_d \in V} d\tau^n \, d\sigma^n_{ab}/d\tau^n; \quad n_c > 0$$

is the total production cross section of $|\, n \,|$ particles, when a particle of sort d does not escape from the domain V.

Averaging $P^n_{ab}(\mathbf{k}_c \mid V,\ d)$ over all channels with a probability $w_n(s \mid V,\ d;\ c) = \sigma^n_{ab}(s \mid V,\ d;\ c) / \sum\limits_m \sigma^m_{ab}(s \mid V,\ d;\ c)$, we arrive at the conditional probability distribution

$$\frac{dw_c}{d^3k_c}(\mathbf{k}_c \mid V,\ d) = \sum_n w_n(s \mid V,\ d;\ c) \, P^n_{ab}(\mathbf{k}_c \mid V,\ d;\ c)$$

$$= \frac{1}{\sigma_{ab \to cdX}(s \mid V,\ d)} \frac{d\sigma_{ab \to cdX}(\mathbf{k}_c \mid V,\ d)}{d^3k_c},$$

where

$$\frac{d\sigma_{ab \to cdX}}{d^3k_c}(\mathbf{k}_c \mid V,\ d) = \sum_n \frac{d\sigma^n_{ab}}{d^3k_c}(\mathbf{k}_c \mid V,\ d)$$

$$= \int\limits_V d^3k_d \frac{d\sigma_{ab \to cdX}}{d^3k_c d^3k_d},$$

$$\sigma_{ab \to cdX}(s \mid V,\ d) = \sum_n \sigma^n_{ab}(s \mid V,\ d;\ c).$$

The relevant inclusive spectrum is apparently

$$f_{ab\to cdX}(s, \mathbf{k}_c \mid V, d) = \sum_n \int_V d^3k_d \frac{d\sigma^n_{ab}}{d^3k_c\, d^3k_d} n_c.$$

By integrating it over $\mathbf{k}_c$, we arrive at the *associated mean multiplicity*

$$\langle n_c\rangle (s \mid V, d) = \sum_n n_c w_n (s \mid V, d).$$

This describes the mean number of the particles of sort c created in those processes where one of the particles of sort d runs through a confined phase space domain V. Obviously, the quantity $\langle n_c\rangle (s \mid V, d) - \delta_{cd}$ describes a mean number of the particles of sort c "accompanying" the particle d which has its momentum range confined within V.

It is not difficult to see that when V becomes the total phase volume we return to the ordinary mean multiplicity. If V degenerates to a point $\mathbf{k}_d$, we have a local associated multiplicity

$$\langle n_c\rangle (s \mid \mathbf{k}_d) = \frac{\sum_n n_c\, d\sigma^n_{ab}/d^3k_d}{\sum_n d\sigma^n_{ab}/d^3k_d} \equiv \sum_n n_c w_n (s \mid \mathbf{k}_d).$$

It can be seen from

$$\int d^3k_d \langle n_c\rangle (s \mid \mathbf{k}_d) \frac{d\sigma_{ab\to cdX}}{d^3k_d} \leqslant \langle n_c\rangle \sigma_{ab\to cX}$$

that the energy evolution of $\langle n_c\rangle (s/\mathbf{k}_d)$ can differ very badly from that of the ordinary mean multiplicity, yet this difference is masked by the cross section $d\sigma_{ab\to cdX}/d^3k_d$.

In the literature one sometimes encounters another definition of the associated mean multiplicity which for the case when $c \neq d$ takes the expression

$$\frac{\int d^3k_c f_{ab\to cdX}(s, \mathbf{k}_c, \mathbf{k}_d)}{f_{ab\to dX}(s, \mathbf{k}_d)} = \sum_n n_c \frac{n_d}{\langle n_d\rangle (s \mid \mathbf{k}_d) + 1} w_n (s \mid \mathbf{k}_d). \tag{2.8}$$

Comparing this expression with the definition above, viz.

$$\langle n_c\rangle (s \mid \mathbf{k}_d) = \sum_n n_c w_n (s \mid \mathbf{k}_d),$$

we see that formula (2.8) yields a correct expression only when $w_n (s \mid \mathbf{k}_d) \sim \delta_{n_d, \langle n_d\rangle (s|\mathbf{k}_d)}$. The approximate equality is, certainly, possible, if the distribution in the multiplicity $w_n (s \mid \mathbf{k}_d)$ is sufficiently closely concentrated around $\langle n_d\rangle (s \mid \mathbf{k}_d)$, i.e. the standard deviation is sufficiently small.

13. *Equivalence of description of multiple production processes in terms of various cross sections.* Both of the above sets of inclusive distributions $d\sigma_{ab\to mCX}/d\tau^m$ and $f_{ab\to mCX}$, are expressed using the same fundamental quantities, i.e. the exclusive cross sections $d\sigma^n_{ab}/d\tau^n$. Whence it follows that the inclusive and exclusive cross sections have to be expressed through each other.

To do this, let us first express the exclusive cross sections $d\sigma^n_{ab}/d\tau^n$ in terms of the differential cross sections $d\sigma_{ab\to mCX}/d\tau^m$. It is easy to see from the definition of $d\sigma_{ab\to mCX}/d\tau^m$ that

$$\frac{d\sigma^n_{ab}}{d\tau^n} = \sum_{m=(0,\ldots,0)}^{(1,\ldots,1)} (-1)^{|m|} \frac{d\sigma_{ab\to(n+m)CX}}{d\tau^n}.$$

Next, from the inclusive spectrum definition

$$f_{ab\to mCX} = \sum_n \frac{n!}{(n-m)!} \frac{d\sigma^n_{ab}}{d\tau^m}$$

and from the relation of $d\sigma^n_{ab}/d\tau^n$ to $d\sigma_{ab\to mCX}/d\tau^m$ just obtained it follows that

$$f_{ab\to mCX} = \sum_l \frac{d\sigma_{ab\to lCX}}{d\tau^m} \frac{l!}{(l-m)!} \prod_{i=1}^{j} \frac{m_i}{l_i}.$$

To express $d\sigma^n_{ab}/d\tau^n$ and $d\sigma_{ab\to mCX}/d\tau^m$ in terms of the inclusive spectra, the following generating functional is suitable:

$$G[h] = \sum_n \prod_{i=1}^{j} \prod_{r=1}^{n_i} h_i(\mathbf{k}_{rc_i})\, d\sigma^n_{ab}/d\tau^n,$$

where $h_i(\mathbf{k}_{rc_i})$ are continuous functions. Whence we have

$$\frac{d\sigma^n_{ab}}{d\tau^n} = \frac{\delta^n G[h]}{n!\,(\delta h)^n}\Big|_{h=0} = \frac{\delta^{|n|} G[h]}{\prod_{i=1}^{j}\prod_{r=1}^{n_i} \delta h_i(\mathbf{k}_{rc_i})} \frac{1}{n!}.$$

Otherwise, it is not difficult to be seen that

$$\frac{\delta^n G[h]}{(\delta h)^n}\Big|_{h=1} = \sum_m \frac{m!}{(m-n)!} \frac{d\sigma^m_{ab}}{d\tau^n} = f_{ab\to nCX}.$$

Whence it follows that $G[h]$ can be presented as

$$G[h] = \sum_n \frac{1}{n!} \int d\tau^n \prod_{i=1}^{j} \prod_{r=1}^{n_i} [h(\mathbf{k}_{rc_i}) - 1]\, f_{ab\to nCX}.$$

Differentiating an approximate number of times by h and letting $h = 0$, we have

$$\frac{d\sigma^n_{ab}}{d\tau^n} = \sum_m \frac{(-1)^{|m-n|}}{(m-n)!} \int d\tau^{m-n} f_{ab\to mCX}.$$

Substituting this expression into the definition of $d\sigma_{ab\to lCX}/d\tau^l$, we obtain an expression for this quantity in terms of inclusive spectra:

$$\frac{d\sigma_{ab\to lCX}}{d\tau^l} = \sum_n \left[\sum_{m=l}^{n} \frac{(-1)^{|n-m|}}{(n-m)!} \right] \int d\tau^{n-l} f_{ab\to nCX}.$$

Thus, any of the three quantities $d\sigma^n_{ab}/d\tau^n$, $d\sigma_{ab\to mCX}/d\tau^m$, $f_{ab\to mCX}$ can be expressed in terms of any of the two remaining ones. This means that full information about any three of the characteristics types is equivalent to full information about the remaining two.

2.3 GENERAL PROPERTIES OF INCLUSIVE DISTRIBUTIONS

2.3.1 Sum Rules

Sum rules are relationships between the various inclusive spectra which follow from the conservation laws for the additive quantities, e.g. 4-momenta, electric, baryonic, and other charges. A method of deriving them is extremely simple. Nevertheless, by virtue of their general nature they have to be necessarily valid within any specific model.

Let us start with the sum rule related to the 4-momentum conservation law which can be written for a definite channel n as

$$k^\nu_a + k^\nu_b = \sum_{i=1}^{j} \sum_{r=1}^{n_i} k^\nu_{rc_i}; \qquad \nu = 0,\ 1,\ 2,\ 3.$$

By multiplying σ^{tot}_{ab} by $k^\nu_a + k^\nu_b$, we can obtain

$$(k^\nu_a + k^\nu_b)\,\sigma^{tot}_{ab} = \frac{1}{\lambda^{1/2}} \sum_n \int d\tau^n\,(k^\nu_a + k^\nu_b)\,(2\pi)^4\,\delta\Big(k_a + k_b - \sum_{ir} k_{rc_i}\Big)\,|\,T^n_{ab}\,|^2 = \frac{1}{\lambda^{1/2}} \sum_{i,\,n} n_i \int d\tau^n\,k^\nu_{1c_i}\,(2\pi)^4$$
$$\times\,\delta\Big(k_a + k_b - \sum_{ir} k_{rc_i}\Big)\,|\,T^n_{ab}\,|^2 = \sum_i \int d^3k_{c_i} k^\nu_{c_i} f_{ab\to c_iX}.$$

The meaning of this identity is self-explaining, i.e. the sum of the mean total momenta $(\sigma_{ab}^{\text{tot}})^{-1} \int d^3k_{c_i} k_{c_i}^{\nu} f_{ab\to c_i X}$ equals the initial momentum.

By increasing the power of the initial momentum, one can obtain sum rules relating to the inclusive spectra of an arbitrarily high order. For example,

$$(k_a^{\mu}+k_b^{\mu})(k_a^{\nu}+k_b^{\nu})\,\sigma_{ab}^{\text{tot}}=\sum_i \Big[\int d^3k_{c_i} k_{c_i}^{\mu} k_{c_i}^{\nu} f_{ab\to c_i X}$$

$$+2\int d^3k_{c_i}\, d^3k'_{c_i}\, k_{c_i}^{\mu} k_{c_i}^{\prime\nu} f_{ab\to c_i c_i X}$$

$$+\sum_{i\neq j} d^3k_{c_i}\, d^3k_{c_j}\, k_{c_i}^{\mu} k_{c_j}^{\nu} f_{ab\to c_i c_j X}\Big]. \quad (2.9)$$

Inclusive spectra themselves can be used in place of σ_{ab}^{tot}. For example,

$$(k_a^{\mu}+k_b^{\mu}-k_c^{\mu})\, f_{ab\to cX} = \sum_d \int d^3k_d k_d^{\mu} f_{ab\to cdX}, \quad c,\ d\in C.$$

Sum rules allow some of the correlations occurring between the particles in a binary inclusive process to be seen even at the kinematics level. From (2.9) in the C.M.S. ($\mathbf{k}_a+\mathbf{k}_b=0$) we can obtain

$$\sum_i \Big[\int d^3k_{c_i}\mathbf{k}_{c_i}^2 f_{ab\to c_i X}+2\int d^3k_{c_i}\, d^3k'_{c_i}\,(\mathbf{k}_{c_i}\cdot\mathbf{k}'_{c_i})\, f_{ab\to c_i c_i X}$$

$$+\sum_{i\neq j}\int d^3k_{c_i}\, d^3k_{c_j}\,(\mathbf{k}_{c_i}\cdot\mathbf{k}_{c_j})\, f_{ab\to c_i c_j X}\Big]=0.$$

This signifies that over the major part of their range, $\mathbf{k}_{c_i}$, $\mathbf{k}'_{c_i}$ and $\mathbf{k}_{c_i}$, $\mathbf{k}_{c_j}$ are directed into different hemispheres. Sometimes a so-called correlation function

$$C_{ij}(s;\ \mathbf{k}_{c_i},\ \mathbf{k}_{c_j})=\frac{f_{ab\to c_i c_j X}(s,\ \mathbf{k}_{c_i},\ \mathbf{k}_{c_j})}{\sigma_{ab\to c_i c_j X}}$$

$$-\frac{f_{ab\to c_i X}(s,\ \mathbf{k}_{c_j})\, f_{ab\to c_j X}(s,\ \mathbf{k}_{c_j})}{\sigma_{ab\to c_i X}\sigma_{ab\to c_j X}}$$

is studied as a measure of the correlation between the particles c_i and c_j which are detected during the course of the inclusive process $a+b\to c_i+c_j+X$. However, the use of such quantities seems to be unjustified for two reasons at any rate:

(1) in a real case with several sorts of particles, there are channels present in $f_{ab\to c_i X}$ and $f_{ab\to c_j X}$ which are not present in $f_{ab\to c_i c_j X}$;

(2) even if we choose identical channels, the fact that a process characteristic does not vanish outside its physical area is hardly attractive and not very understandable.

Suppose Q is an electric charge, then clearly

$$Q_a + Q_b = \sum_j \langle n_j \rangle Q_j.$$

The meaning of this sum rule is clear-cut, the sum of the average charges carried away by the particles is equal to the initial charge.

Analogous statements are true for other additive charges as well, for example, baryonic, leptonic, etc.

2.3.2 Analytic Properties of the Cross Sections of Multiple Production Processes with Respect to the Cosine of the Emission Angle of the Detected Particle

In the previous section we treated general properties of inclusive cross sections (the sum rules) which are artifacts of only the additive conservation laws. The goal of the present section is to establish some more properties of the cross sections which follow from principles such as microcausality, spectrality, and unitarity.

Discussion of the analytic properties of physical quantities with respect to some variable is, as a rule, an extrapolation from a physical domain (i.e. from one where a quantity is measurable) to a domain of values which have no direct physical meaning, for example the domain of complex angular values. Why is this necessary?

In addition to finding information that may be potentially interesting from various points of view, there is quite a practical side. Analytic functions known are to be exceedingly "rigid" objects in that their behavior in different domains is highly correlated. In the case of elastic scattering this is manifested in the behavior of the forward scattering amplitude being controlled mainly by the position of the nearest singularity in the plane of complex scattering angles. It is natural to expect something similar in the case of multiple production processes.

In quantum field theory the analytic properties of amplitudes are closely related to the causality principle, and in order to state it, we have to deal with local quantities that are like field operators. With Bogolyubov's formulation [10], the one we are going to use, the role of local quantities is played by "radiation operators" $[\delta^n S/\delta\varphi_1(x_1) \ldots \delta\varphi_n(x_n)]\, S^+$ and by their functional derivatives.

Consider the amplitude of a process $a + b \to nC$:

$$\langle nC \,|S|\, a,\, b\rangle = i\,(2\pi)^4\,\delta\,(k_a + k_b - \sum_{il} k_{lc_i})\, T^n_{ab},$$

where we are limited to the case where the number of final particles is more than two. Using Bogolyubov's reduction techniques, we obtain

$$\langle nC|S|a,\, b\rangle = -i \int dx\; e^{i\frac{k_a-k_b}{2}x} \left\langle nC \left| \frac{\delta J_a(-x/2)}{\delta\varphi_b(x/2)} \right| 0 \right\rangle \times (2\pi)^4\, \delta\Big(k_a+k_b-\sum_{il} k_{lc_i}\Big),$$

i.e.

$$T^n_{ab} = -\int dx\, e^{i\frac{k_a-k_b}{2}x} \left\langle nC \left| \frac{\delta J_a(-x/2)}{\delta\varphi_b(x/2)} \right| 0 \right\rangle,$$

where $J_a(x) = i\,[\delta S/\delta\varphi_a(x)]\, S^+$ is a source operator for the particles of sort a.

The conditions of microcausality,

$$\delta J_a\left(-\frac{x}{2}\right)\Big/\delta\varphi_b\left(\frac{x}{2}\right) = 0, \quad x^0 < |\mathbf{x}|,$$

and spectrality,

$$\int dx\; e^{i\frac{k_a-k_b}{2}} \langle nC|[J_a(-x/2),\; J_b(x/2)]|0\rangle = 0,$$

outside the domains

$$\Big[\Big(\sum_{il} k_{lc_i}\Big) \pm \frac{1}{2}(k_a-k_b)\Big]^2 > 0,$$

$$\Big(\sum_{il} k^0_{lc_i}\Big) \pm \frac{1}{2}(k^0_a-k^0_b) > 0,$$

mean that the Jost-Lehmann-Dyson representation can be used. This is obtained for any amplitude possessing the following property of disappearance:

$$\langle nC|T|ab\rangle = \int d^4u\, d\lambda^2 \frac{\psi(u,\, \lambda^2;\, \mathbf{K}_n)}{\left[\left(\frac{k_a-k_b}{2}\right)-u\right]^2-\lambda^2}.$$

A support of the spectral function ψ in u and λ^2 is confined within the domain

$$\frac{k^0_a+k^0_b}{2} \pm u^0 > \left|\frac{\mathbf{k}_a+\mathbf{k}_b}{2} \pm \mathbf{u}\right|,$$

$$\lambda \geqslant \max\left[0,\; m_1-\sqrt{\left(\frac{k_a+k_b}{2}-u\right)^2},\; m_2-\sqrt{\left(\frac{k_a+k_b}{2}+u\right)^2}\right], \qquad (2.10)$$

where $m_{1,2}$ are masses of the lowest states $|\bar{n}_{1,2}\rangle$ for which

$$\langle nC|J_a|\bar{n}_1\rangle\langle\bar{n}_1|J_b|0\rangle\neq 0,$$
$$\langle nC|J_b|\bar{n}_2\rangle\langle\bar{n}_2|J_b|0\rangle\neq 0.$$

It can be seen from the expression for $\langle nC \mid T \mid ab\rangle$ that once the direction of initial particles is given by a unit vector $\mathbf{n} = \mathbf{k}_a / \mid \mathbf{k}_b \mid$ in the C.M.S. ($\mathbf{k}_a + \mathbf{k}_b = 0$) the vector $\mathbf{n}$ only enters the denominator in a reference frame relative to the configuration of the final momenta $\mathbf{K}_n$. This allows the analytic properties of the amplitude $\langle nC \mid T \mid ab\rangle$ to be studied with respect to the vector $\mathbf{n}$, with the analyticity domain clearly being determined from the spectrality condition expressed in the form of a support of the spectral function ψ in u and λ^2. The fact that the boundaries of the support of ψ (see Eq. (2.10)) depend only on the total energy $E_a+E_b=\sqrt{s}$ in C.M.S. is important. It follows from this that the analyticity domain in $\mathbf{n}$, too, will only depend on s, rather than on the remaining variables that specify the final state $\langle nC \mid$. The general derivation of the analytic properties of the amplitude $\langle cn \mid T \mid ab\rangle$ with respect to the vector $\mathbf{n}$ [11] can be found in the Appendix. Yet here we shall limit ourselves to a simple derivation of the analytic properties with respect to only one of the two independent variables that define the vector $\mathbf{n}$ in a spherical parametrization:

$$\mathbf{n} = (\cos\varphi\sin\theta,\ \sin\varphi\sin\theta,\ \cos\theta),$$

viz. with respect to $z = \cos\theta$. The method that can do this dates back to the works [1, 12].

If $\mid n \mid$ is the number of the final particles, then the amplitude $\langle nC \mid T \mid ab\rangle$ depends on $3 \mid n \mid - 4$ independent variables. In a frame of reference with the axis Ox defined by the direction of a momentum $\mathbf{k}_c \in \mathbf{k}_n$ and a momentum $\mathbf{k}_a$ lying in the plane xOy, we choose the following independent variables: $s = (k_a + k_b)^2$; θ, the angle between the vectors $\mathbf{k}_a$ and $\mathbf{k}_c$, $(\mid n \mid - 1)$ independent final particles energies; and $2(\mid n \mid - 2) - 1 = 2 \mid n \mid - 5$ spherical angles α_i, β_i (one of these angles can be expressed in terms of the others due to the 4-momentum conservation law).

The Jost-Lehmann-Dyson representation can conveniently be reduced to the following form:

$$\langle nC|T|ab\rangle = T^n_{ab}(s,\theta,\ \ldots,\ \alpha_i,\ \beta_i,\ E_i,\ \ldots)$$
$$= \int\limits_{x_L(s)}^{\infty} dx \int\limits_0^{2\pi} d\alpha\, \frac{\Phi(x,\alpha;\ s,\ \ldots,\ \alpha_i,\ \beta_i,\ E_i,\ \ldots)}{x-\cos(\theta-\alpha)},$$

where

$$\Phi(x, \alpha; s, \ldots, \alpha_i, \beta_i, E_i, \ldots)$$
$$= \int du^0 \int d|\mathbf{u}|\,|\mathbf{u}| \int \frac{d\cos\beta}{2\,|\mathbf{k}_a|\sin\beta} \int d\lambda^2$$
$$\times \delta\left(x - \frac{(\mathbf{u}^2 + \mathbf{k}_a^2 + \lambda^2) - (u^0 - (k_a^0 - k_b^0)/2)^2}{2\,|\mathbf{u}|\,|\mathbf{k}_a|\sin\beta}\right) \psi(u^0, \mathbf{u}, \lambda^2, \ldots),$$

α and β are the azimuthal and polar angles of the vector $\mathbf{u}$ and the integration domain over x is defined by the form of the domain, where $\psi \neq 0$:

$$x_L(s) = \sqrt{1 + \frac{(m_1^2 - m_b^2)(m_2^2 - m_a^2)}{\mathbf{k}_a^2 (s - (m_1 - m_2)^2)}}.$$

An expression in the denominator of the integral representation contains $\sin\theta$. If we are interested in analytic properties in $\cos\theta$, $\sin\theta$ results in singularities at $\cos\theta = \pm 1$, which occur for spherical coordinates at $\theta = 0, \pi$. In this case, for example, the plane xOy is not, as yet, fixed by the vectors $\mathbf{k}_a$ and $\mathbf{k}_c$ but the singularity can be isolated in an explicit form.

In fact, since the denominator in the integral representation has the form

$$x - \cos(\theta - \alpha) = x - \cos\theta\cos\alpha - \sin\theta\sin\alpha,$$

then, multiplying and dividing the integrand by

$$x - \cos\theta\cos\alpha + \sin\theta\sin\alpha$$

will yield

$$T_{ab}^n = \int\limits_{x_L(s)}^{\infty} dx \int\limits_0^{2\pi} d\alpha\, \Phi \frac{x - \cos\theta\cos\alpha + \sin\theta\sin\alpha}{(x - \cos\theta\cos\alpha)^2 - \sin^2\alpha + \cos^2\theta\sin^2\alpha}$$
$$\equiv T_{ab,1}^n(s, \cos\theta, \ldots) + \sin\theta\, T_{ab,2}^n(s, \cos\theta, \ldots).$$

Now the functions $T_{ab,1,2}^n$ do not contain odd powers of $\sin\theta$ and their singularities with respect to $\cos\theta$ are defined minimizing the domain with the vanishing denominator in the integral representation over x and α. This takes place at

$$\cos\theta = x\cos\alpha \pm i\sqrt{x^2 - 1}\sin\alpha.$$

It is easy to see that, while integrating over α, we run over an ellipse in the complex plane of $\cos\theta$ which has foci at ± 1 and its semi-major axis x. A minimum ellipse can clearly be obtained at $x = x_L(s)$. Thus, the amplitudes $T_{ab,1}^n$, $T_{ab,2}^n$ are analytic in $\cos\theta$ within the domain confined by a Lehmann ellipse with the foci at ± 1 and a semi-major axis $x_L(s)$. Yet the full amplitude $T_{ab}^n(s, \cos\theta, \ldots)$ will have additional branch points at $\cos\theta = \pm 1$

as well. To fix the sign of $\sin\theta = \sqrt{1-\cos^2\theta}$ the cuts $(-\infty, -1]$ and $[1, +\infty)$ must be drawn.

Now let us pass on to the production cross section of a particle c at an angle θ with respect to the direction of motion of initial particles. This is a partially integrated cross section having, according to Sec. 2.2, the form

$$\frac{d\sigma_{ab}^{n}}{d\cos\theta} = \int \frac{\mathbf{k}_c^2 d|\mathbf{k}_c|}{2E_c(2\pi)^2} \frac{d\sigma_{ab}^{n}}{d^3k_c}$$

$$= \frac{1}{\lambda^{1/2}(s)} \int \frac{\mathbf{k}_c^2 d|\mathbf{k}_c|}{2E_c(2\pi)^2} d\tau^{n'} (2\pi)^4 \delta\Big(k_a + k_b - k_c$$

$$- \sum_{i=1}^{j} \sum_{l=1}^{n_i'} k_{lc_i}\Big) |T_{ab}^{n}|^2 \equiv \int d\Gamma_n^c |T_{ab}^n|^2,$$

$$n' = n - (0, 0, \ldots, 1_c, \ldots, 0).$$

Now note that owing to an invariance relative to rotations in three dimensions, we have

$$\frac{d\sigma_{ab}^{n}}{d\cos\theta}(\theta, s) = \frac{d\sigma_{ab}^{n}}{d\cos\theta}(2\pi - \theta, s),$$

whence it follows that

$$\frac{d\sigma_{ab}^{n}}{d\cos\theta}(\theta, s) = \frac{1}{2}\left[\frac{d\sigma_{ab}^{n}}{d\cos\theta}(\theta, s) + \frac{d\sigma_{ab}^{n}}{d\cos\theta}(2\pi-\theta, s)\right]$$

$$= \frac{1}{2}\int d\Gamma_n^c [|T_{ab}^n(s, \theta, \ldots)|^2 + |T_{ab}^n(s, 2\pi-\theta, \ldots)|^2]$$

$$= \int d\Gamma_n^c [|T_{ab,1}^n(s, \cos\theta, \ldots)|^2$$

$$+ \sin^2\theta\, |T_{ab,2}^n(s, \cos\theta, \ldots)|^2].$$

Since $T_{ab,1,2}^n$ depends only on the even powers of $\sin\theta$ and does not change in going from $\theta \to 2\pi - \theta$, terms of the form $\sin\theta$ $(\overset{*}{T}{}_{ab,1}^{n} T_{ab,2}^{n} + T_{ab,1}^{n} \overset{*}{T}{}_{ab,2}^{n})$ cancel out due to the sign difference between $\sin\theta$ and $\sin(2\pi - \theta)$. Thus, we see that the cross section is analytically continued in the Lehmann ellipse and does not contain any physical singularities at $\theta = 0, \pi$.

Since the sum of a finite number of the terms is analytic in the domain and the members of the sum are also analytic in this domain, it follows that for any fixed s, when the number of channels is limited, the inclusive differential cross sections

$$\frac{d\sigma_{ab\to cX}}{d\cos\theta} = \int \frac{\mathbf{k}_c^2 d|\mathbf{k}_c|}{8\pi^2 E_c} \frac{d\sigma_{ab\to cX}}{d^3k_c} = \sum_n \frac{d\sigma_{ab}^n}{d\cos\theta}$$

and the inclusive spectra

$$\hat{f}_{ab\to cX}(s,\ \cos\theta) = \int \frac{\mathbf{k}_c^2 d|\mathbf{k}_c|}{8\pi^2 E_c} f_{ab\to cX} = \sum_n n_c \frac{d\sigma_{ab}^n}{d\cos\theta}$$

are functions which are analytic in $\cos\theta$ and regular within the domain confined by the ellipse whose foci are at $\cos\theta = \pm 1$ and whose semi-major axis is $x_L(s)$.

2.3.3 Analytical Properties in Two Angular Variables [13]

In Sec. 2.3.2 we considered the analytical properties of the amplitudes and cross sections of multiple processes with respect to the cosine of an angle between the initial and detected particles. As is seen from the drawings, θ is the angle defining the direction of the initial momentum $\mathbf{k}_a$ relative to a rigidly fixed system of final momenta. The second, azimuthal angle specifying the direction of $\mathbf{k}_a$ is related to the final particles through the choice of axes.

Now we want to investigate the analytical properties in both angles specifying the direction of $\mathbf{k}_a$. To do this, it is convenient for the coordinate axes to be somewhat altered. The axis z now points along the momentum vector $\mathbf{k}_c$ and the plane xOz determining the origin of reference of the azimuthal angles is built up from the vectors $\mathbf{k}_c$ and $\mathbf{k}_d$, where d is any, yet fixed, particle from the final state. The angles $(\theta,\ \varphi)$ describe the direction of the vector $\mathbf{n} = \mathbf{k}_a / |\mathbf{k}_a|$. An explicit form for the remaining independent variables is now inessential for our purpose and is denoted as a whole by ξ_n.

The Jost-Lehmann-Dyson representation for the amplitude of the process

$$a + b \to nC$$

will be rewritten in the form

$$T_{ab}^n(s,\ \mathbf{n},\ \ldots) = \int\limits_{x_L(s)}^{\infty} dx \int d\mathbf{e} \frac{\Phi(x,\ \mathbf{e};\ s,\ \ldots)}{x - \mathbf{e}\cdot\mathbf{n}},$$

where $x_L(s)$ is the same as in Sec. 2.2, $\mathbf{e} = \mathbf{u}/|\mathbf{u}|$ and

$$\Phi(x, \mathbf{e}; s, \ldots) = \frac{1}{2|\mathbf{k}_a|} \int d\mathbf{u}_0 d|\mathbf{u}||\mathbf{u}|\,\psi(u, \lambda^2; s, k_c, k_d, \ldots)$$

$$\times\,\delta\left(x - \frac{\mathbf{k}_a^2 + \mathbf{u}^2 + \lambda^2 - \left(\frac{m_a^2 - m_b^2}{\sqrt{s}} - u_0\right)^2}{2|\mathbf{k}_a|\,|\mathbf{u}|}\right).$$

As is shown in the Appendix, $T_{ab}^n(s,\ \mathbf{n},\ \ldots)$ is the analytic function of a complex vector $\mathbf{n}$ within a certain domain on a complex

three-dimensional sphere whose dimensions are controlled by the value of $x_L(s)$. Using a spherical parametrization

$$\mathbf{n} = (\cos\varphi \sin\theta,\ \sin\varphi \sin\theta,\ \cos\theta),$$

this domain has the form

$$\begin{gathered}(|1+z|\,|w|^{-1}+|1-z|\,|w|)(|1+z|\,|w| \\ +|1-z|\,|w|^{-1}) < +\ 4x_L^2(s), \\ z \,\bar{\in}\, (-x_L(s),\ -1] \cup [1,\ x_L(s)),\end{gathered} \tag{2.11}$$

where $z = \cos\theta$ and $w = e^{i\varphi}$. The appearance of cuts from $z = \pm 1$, as indicated above, is related to the fact that a unique parametrization that is regular everywhere cannot be chosen for the sphere. For the given case of parametrization in terms of the spherical angles θ and φ there is a singularity at $\theta = 0, \pi$.

Provided the angle φ is real ($|w| = 1$), we obtain the result in Sec. 2.3.2, i.e. the analyticity inside the Lehmann ellipse:

$$|1-z| + |1+z| < 2x_L(s)$$

with cuts $(-x_L(s),\ -1] \cup [1,\ x_L(s))$. By solving inequality (2.11) with respect to $|w|$, we can obtain a circular domain in the w-plane with the radii depending on z:

$$\begin{gathered}R_-(z) < |w| < R_+(z), \\ R_\pm(z) \\ = \frac{[4x_L^2-(|1+z|-|1-z|)^2]^{1/2} \pm [4x_L^2-(|1+z|+|1-z|)^2]^{1/2}}{2\,|1-z^2|^{1/2}}.\end{gathered}$$

For the inequality to be satisfied, z must belong to the Lehmann ellipse.

It is evident that

$$R_+R_- = 1$$

and

$$R_- \leqslant 1 \leqslant R_+.$$

The ring has its minimum size at $z = \cos\theta = 0$, i.e. at $\theta = \pi/2$, when

$$R_\pm(s) = x_L(s) \pm \sqrt{x_L^2(s)-1}.$$

The conjugate amplitude $\overset{*}{T}{}^n_{ab}(s, \mathbf{n}, \ldots)$ obviously continues in $\mathbf{n}$ into the same domain as does $T^n_{ab}(s, \mathbf{n}, \ldots)$.

Whence we obtain that the differential cross section,

$$\begin{gathered}\frac{d\sigma^n_{ab}}{d\cos\theta\, d\varphi} = \int d^3k_c\, d^3k_d\, \frac{d\sigma^n_{ab}}{d^3k_c\, d^3k_d}\, \delta\left(\cos\theta - \frac{\mathbf{k}_a\mathbf{k}_c}{|\mathbf{k}_a|\,|\mathbf{k}_c|}\right) \\ \times \delta\left(\varphi - \arccos\frac{\mathbf{k}_c^2(\mathbf{k}_a\mathbf{k}_d)-(\mathbf{k}_c\mathbf{k}_d)(\mathbf{k}_c\mathbf{k}_a)}{\sqrt{[\mathbf{k}_c^2\mathbf{k}_d^2-(\mathbf{k}_c\mathbf{k}_d)^2]\,[\mathbf{k}_a^2\mathbf{k}_c^2-(\mathbf{k}_a\mathbf{k}_c)^2]}}\right),\end{gathered}$$

analytically continues into domain (2.11) as a function of two complex variables $x = \cos\theta$ and $w = e^{i\varphi}$.

This, in turn, leads to the inclusive differential cross section being analytic in domain (2.11),

$$\frac{d\sigma_{ab\to c\, dX}}{d\cos\theta\, d\varphi} = \sum_n \frac{d\sigma^n_{ab}}{d\cos\theta\, d\varphi},$$

and the inclusive binary spectrum

$$\hat{f}_{ab\to c\, dX}(s,\ \theta,\ \varphi) = \sum_n n_c n_d \frac{d\sigma^n_{ab}}{d\cos\theta\, d\varphi},$$

which describe the inclusive process

$$a + b \to c + d + X.$$

We specified the initial momentum $\mathbf{k}_a$ above in terms of the angles θ and φ. The quantities θ and φ can also be regarded as the production angle of the particle c and the angle between the planes $[\mathbf{k}_a\mathbf{k}_c]$ and $[\mathbf{k}_c\mathbf{k}_d]$, these angles being considered here as characteristics of the detected particles.

2.3.4 A Unitarity Condition and the Analytical Properties of the Characteristics of Multiple Production Processes in the Angular Variables Θ and φ [13, 14]

Up to now we have used only the "linear" general principles, i.e. microcausality and spectrality, when studying the analytic properties of the amplitudes and cross sections of multiple processes in angular variables. These are expressed in the form of some linear conditions on the amplitudes of physical processes, i.e. not containing higher powers of the amplitudes in a nontrivial fashion. The unitarity condition,

$$SS^+ = 1, \tag{2.12}$$

refers to the squares of the amplitude moduli of physical processes, i.e. it is a nonlinear condition and must, generally, bring about new properties of the amplitudes. Consider operator equality (2.12) for matrix elements between the two-particle states $\langle a',\ b' \mid$ and $\mid a,\ b\rangle$. By putting in a complete system of intermediate states and isolating the connected parts, we obtain

$$\frac{1}{i}\langle a', b' | T - T^+ | a, b\rangle = \sum_n d\tau^n (2\pi)^4$$
$$\times\, \delta\Big(k_a + k_b - \sum_{i=1}^{j} \sum_{l=1}^{n_i} k_{lc_i}\Big) \langle a',\ b' | T^+ | n\rangle \langle n | T | a,\ b\rangle. \tag{2.13}$$

It follows from the invariance with respect to the PT-transformation that the left-hand side of this equation is equal to $2 \operatorname{Im} \langle a', b' \mid T \mid a, b \rangle$, i.e. to the imaginary part of the elastic scattering amplitude.

Thus, the unitarity condition leads to a correlation between the properties of elastic and inelastic amplitudes. As one takes other matrix elements of the unitarity condition, then, clearly, each of the relations obtained will deliver more new links between the amplitudes of various processes. The concrete resolution of these bonds is extremely complicated. So far only the corollaries from the two-particle unitarity condition (2.13) have been sufficiently well studied and some first steps are being made in studying the unitarity condition for the $3 \to 3$ process (Sec. 2.5).

Here we want to see what the result of the unitarity condition is in terms of the analytic properties of the amplitudes and cross sections of inelastic processes.

As has been shown by Martin [15], the imaginary part of the elastic scattering amplitude $\operatorname{Im} \langle a', b' \mid T \mid a, b \rangle \equiv \operatorname{Im} T_2(s, \mathbf{n} \cdot \mathbf{n}')$, where $\mathbf{n} = \mathbf{k}_a / \mid \mathbf{k}_a \mid$, $\mathbf{n}' = \mathbf{k}'_a / \mid \mathbf{k}_a \mid$, is analytically continued in $\cos\theta = \mathbf{n} \cdot \mathbf{n}'$ into an ellipse with foci at $\cos\theta = \pm 1$ and semi-major axis $x_M(s) = 1 + 2m_\pi^2 / \mathbf{k}_a^2 \cong 1 + 8m_\pi^2/s$, $s \to \infty$. This means that the nearest singularity is in the t-plane at the point $4m_\pi^2$, i.e. indicating an exchange by two π-mesons.

Expand the elastic scattering amplitude into a series in Legendre polynomials (the normalizing coefficients here and later are chosen for convenience only), viz.

$$T_2(s, \cos\theta) = \frac{8\pi \sqrt{s}}{|\mathbf{k}_a|} \sum_l (2l+1) f_l(s) P_l(\cos\theta),$$

and the amplitudes of the process $ab \to nC$, into a series in spherical functions $D_{m0}(\mathbf{n})$:

$$T^n_{ab}(s, \mathbf{n}, \ldots) = \left(\frac{2\sqrt{s}}{|\mathbf{k}_a|} \right)^{1/2} \sum_{l,\, m} (2l+1) T^{n;\, lm}_{ab}(s, \ldots) D^l_{m0}(\mathbf{n}). \qquad (2.14)$$

In terms of partial amplitudes, the unitarity condition takes the form

$$\operatorname{Im} f_l(s) = |f_l(s)|^2 + \sum_n \sum_{m=-l}^{l} \int d\xi_n |T^{n;\, lm}_{ab}(s, \xi_n)|^2. \qquad (2.15)$$

From the analyticity of $\operatorname{Im} T_2(s, z)$ in the Martin ellipse an inequality follows

$$\operatorname{Im} f_l(s) \leqslant \frac{s^N e^{-l \ln (x_M + \sqrt{x_M^2 - 1})}}{\sqrt{2l+1}}, \qquad (2.16)$$

where s^N is the upper bound for the evolution of $\operatorname{Im} T_2(s, z)$ in s in the analyticity domain.

Equation (2.15) results in a limitation being imposed on the partial amplitudes of inelastic processes:

$$\sum_n \sum_{m=-l}^{l} \int d\xi_n \, |T_{ab}^{n,\, lm}(s, \xi_n)|^2 < \operatorname{Im} f_l(s). \tag{2.17}$$

It is impossible to deduce any definite conclusions therefrom that are related to the behavior of $T_{ab}^{n,\, lm}(s, \xi_n)$, since the partial amplitudes enter inequality (2.17) having been integrated over ξ_n. Therefore, we shall study the analytic properties of the cross sections rather than those of the amplitudes.

The cross section $d\sigma_{ab}^n/d\cos\theta\, d\varphi$ can be represented in the form

$$\frac{d\sigma_{ab}^n}{d\cos\theta\, d\varphi} = \frac{1}{\mathbf{k}_a^2} \sum_{lm} \sum_{l'm'} (2l+1)(2l'+1) \times \overset{*}{D}{}_{m0}^{l}(\mathbf{n})\, D_{m'0}^{l'}(\mathbf{n}) C_{mm'}^{ll'}(n, s), \tag{2.18}$$

where

$$C_{mm'}^{ll'}(n, s) = \int d\xi_n \overset{*}{T}{}_{ab}^{n,\, lm}\, T_{ab}^{n,\, l'm'}.$$

Where does this series converge? It has been shown above (see Sec. 2.3.2) that the amplitudes of the process $ab \to nC$ in $z = \cos\theta$ and $w = e^{i\varphi}$ are analytic in the domain

$$(|1+z|\,|w| + |1-z|\,|w|^{-1})(|1+z|\,|w|^{-1} + |1-z|\,|w|) < 4x_L^2(s),$$
$$z \bar{\in} (-x_L(s), -1] \cup [1, x_L(s)).$$

One can show that this domain contains the relevant domain for convergence (2.14), namely

$$\frac{1}{2}(|w| + |w|^{-1})(|1+z| + |1-z|) < 2x_L(s), \tag{2.19}$$
$$z \bar{\in} (-x_L(s), -1] \cup [1, x_L(s)).$$

Thus, $d\sigma_{ab}^n/d\cos\theta\, d\varphi$ can be represented in the form of (2.18) in domain (2.19), where the series converges uniformly and absolutely. However, now we have some additional information about the coefficients $C_{mm'}^{ll'}(n, s)$, obtained once we take into account the unitarity condition (2.17).

From the Bunyakovsky-Schwarz inequality we have

$$|C_{mm'}^{ll'}(n, s)| \leqslant \sqrt{C_{mm}^{ll}(n, s)\, C_{m'm'}^{l'l'}(n, s)} \leqslant \sqrt{\operatorname{Im} f_l(s) \operatorname{Im} f_{l'}(s)}.$$

This inequality, along with inequality (2.16) for $\operatorname{Im} f_l(s)$ and that for $D^l_{m0}(\theta, \varphi)$ at the physical values of φ and $\theta = \alpha + i\beta$:

$$|D^l_{m0}(\theta, \varphi)| \leqslant [\cosh\beta + \sqrt{\cosh^2\beta - 1}]^l$$

can be used to evaluate the convergence domain of expansion (2.18):

$$\frac{d\sigma^n_{ab}}{d\cos\theta\, d\varphi} \leqslant \frac{1}{\mathbf{k}^2_a} \sum_{l,\, l'} (2l+1)(2l'+1)[\cosh\beta + \sqrt{\cosh^2\beta - 1}]^{l+l'} \sum_{m,\, m'} \sqrt{C^{ll}_{mm}(n, s)\, C^{l'l'}_{m'm'}(n, s)}$$

$$\leqslant \frac{1}{\mathbf{k}^2_a}\Big[\sum_l (2l+1)^{3/2}(\cosh\beta + \sqrt{\cosh^2\beta - 1})^l \sqrt{\operatorname{Im} f_l(s)}\Big]^2$$

$$\leqslant \frac{s^N}{\mathbf{k}^2_a}\Big[\sum_l (2l+1)^{5/4}\Big(\frac{\cosh\beta + \sqrt{\cosh^2\beta - 1}}{\sqrt{x_M(s) + \sqrt{x^2_M(s) - 1}}}\Big)^l\Big]^2.$$

Whence it follows immediately that series (2.18) converges at the physical values of φ within an ellipse

$$|1+z| + |1-z| < 2x_0(s) = 2\sqrt{\frac{1 + x_M(s)}{2}}, \qquad (2.20)$$

excluding the cuts $(-x_0(s), -1]$ and $[1, x_0(s))$. In terms of the spherical functions $D^l_{m0}(\theta, \varphi) = e^{im\varphi} d^l_{m0}(\theta)$, where $d^l_{m0}(\theta)$ are the Wigner functions, the origin of the branching at $z = \cos\theta = \pm 1$ is associated with the presence of the branching in $d^l_{m0}(\theta)$, which can be represented in the form

$$d^l_{m0}(\theta) = (1 - \cos^2\theta)^{|m|/2} R^l_m(\theta),$$

$R^l_m(\theta)$ being regular at $\cos\theta = \pm 1$. Integrating over φ results in $e^{i(m-m')\varphi}$ yielding $\delta_{mm'}$ and the disappearance of the branching. Whence it follows that the cross section $d\sigma^n_{ab}/d\cos\theta$ continues analytically into the domain (2.20), but now without the cuts from ± 1.

Thus, the unitarity condition substantially enlarges the analyticity domain

$$x_0(s) \simeq 1 + \frac{2m^2_\pi}{s},$$

whereas

$$x_L(s) \simeq 1 + \frac{m^4}{s^2}.$$

It is evident that the inclusive cross sections

$$\frac{d\sigma_{ab\to cX}}{d\cos\theta} = \sum_n \frac{d\sigma^n_{ab}}{d\cos\theta} \quad\text{and}\quad \frac{d\sigma_{ab\to cdX}}{d\cos\theta\, d\varphi} = \sum_n \frac{d\sigma^n_{ab}}{d\cos\theta\, d\varphi}$$

and the inclusive spectra

$$f_{ab\to cX}(s, \cos\theta) = \sum_n n_c \frac{d\sigma^n_{ab}}{d\cos\theta},$$

$$f_{ab\to c\,dX}(s, \cos\theta, \varphi) = \sum_n n_c n_d \frac{d\sigma^n_{ab}}{d\cos\theta\, d\varphi}$$

possess the same analytic properties, as do the relevant exclusive cross sections $d\sigma^n_{ab}/d\cos\theta$ and $d\sigma^n_{ab}/d\cos\theta\, d\varphi$.

2.4 LIMITATIONS ON THE BEHAVIOR OF HIGH ENERGY INCLUSIVE DISTRIBUTIONS

In the preceding section we considered the general properties of inclusive distributions which we obtained on the basis of theoretical fundamentals. However, these properties *per se* (for example, analytical angular properties) cannot be correlated with experimental data.

Therefore, we shall in this section be concerned with deducing from these generalities corollaries which are closely related to the behavior of inclusive distributions in various subdomains of a physical domain. It should be borne in mind that starting from the small number of initial assumptions we have made within the framework of the theoretical fundamentals, we cannot hope to acquire highly detailed information on the behavior of physical characteristics. Indeed, even remaining within the framework of these fundamentals, we still have a significant number of amplitude properties to select and so a wide spectrum of very differing interaction mechanism is possible.

The results we will obtain in the present section will have the form of inequalities restricting (but not contradicting our adopted general principles) the possible evolution of inclusive distributions with energy.

Certainly, with no dynamics of interaction specified, we cannot say anything about why, for example, scaling is violated in a central domain. However, we can answer questions about the degree of the violation.

On the other hand, the value of these sorts of restrictions is their independence of the particular choice of the interaction mechanism and they must be obeyed in any model and so comparison with these general results is a reliable method of checking the features of some model approach. In addition, it is conceivable that the bounds obtained from the theoretical fundamentals can contradict the experimental data, in which case we will be forced to revise the funda-

mentals, replacing them by new ones or limiting their applicability. As is demonstrated by history of physics, such a revision always leads to a greater insight into nature.

2.4.1 Rigorous Bounds on Inclusive Distributions [1, 8, 13, 16]

Before making our attack on the derivation of those limitations on inclusive spectra that follow from their analytic properties, let us see what can be gained from the sum rules.

It has been shown in Sec. 2.3 that σ_{ab}^{tot} and the single-particle inclusive spectra are related via the following sum rule:

$$\sqrt{s}\,\sigma_{ab}^{\text{tot}}(s) = \frac{1}{2(2\pi)^3}\sum_c \int d\mathbf{k}_c f_{ab\to cX}(s, \mathbf{k}_c),$$

where $f_{ab\to cX}$ is the inclusive spectrum of the process

$$a + b \to c + X.$$

In terms of the variables $x = 2k_{\parallel}/\sqrt{s}$ and $\mathbf{k}_{\perp}$, where the components $k_{\parallel}$, $\mathbf{k}_{\perp}$ are defined with respect to the direction of motion of the initial particles in their center-of-inertia frame, we have

$$\sum_c \int dx\, d\mathbf{k}_{\perp} f_{ab\to cX}(s, x, \mathbf{k}_{\perp}) = 32\pi^3\sigma_{ab}^{\text{tot}}(s).$$

It follows from the Froissart bound,

$$\sigma_{ab}^{\text{tot}}(s) \leqslant \frac{\pi}{m_\pi^2}\ln^2\frac{s}{s_0}; \quad s \gg s_0,$$

that

$$\int dx\, d\mathbf{k}_{\perp} f_{ab\to cX}(s, x, \mathbf{k}_{\perp}) \leqslant \frac{32\pi^4}{m_\pi^2}\ln^2\frac{s}{s_0}.$$

Likewise, one can obtain the following limitation:

$$\int dx_c\, d\mathbf{k}_{\perp c}\, dx_d\, d\mathbf{k}_{\perp d} f_{ab\to cdX} \leqslant \frac{4(2\pi)^7}{m_\pi^2}\ln^2\frac{s}{s_0}$$

etc. for the binary inclusive spectrum related to the process $a + b \to c + d + X$.

The resultant inequalities indicate that the inclusive spectra cannot uniformly increase faster than $\ln^2 s$ inside the domain of the variables describing the motion of the detected particles. However, in some of the domains of the momenta of the detected particles a fast increase of spectra does not contradict the above inequalities, provided that the dimensions of the cited domains are reduced sufficiently quickly with energy. To study such situations, an integration over at least some of the variables associated with the detected

particles must be eliminated. Here also the analytical properties of inclusive distributions obtained in Sec. 2.3 come to our aid. Since the analytical properties were obtained with respect to angular variables, all the following will be related to angular distributions.

We begin with a single particle angular distribution in the exclusive process $ab \to Cn$:

$$\Phi_n(s, \cos\theta) = \frac{1}{\sigma_{ab}^n(s)} \frac{d\sigma_{ab}^n}{d\cos\theta}.$$

Then one can obtain bounds on the inclusive distributions from the bounds on Φ_n by summing them (with relevant weights) over all channels. As shown above, Φ_n continues analytically into an ellipse with foci at $\cos\theta = \pm 1$ and a semi-major axis

$$x_0(s) \simeq 1 + \frac{2m_\pi^2}{s}.$$

In this domain a series in Legendre polynomials converges thus,

$$\Phi_n(s, \cos\theta) = \sum_l (2l+1)\, a_l(s)\, P_l(\cos\theta),$$

with $|a_l(s)| \leqslant a_0 = 1/2$, since

$$a_l(s) = \frac{1}{2} \int_{-1}^{1} dz\, P_l(z)\, \Phi_n(s, z).$$

Assuming that $\Phi_n(s, z)$ is bounded polynomially in s inside the analyticity domain, we can show that (at $l \gg 1$)

$$|a_l(s)| \leqslant \frac{s^N \exp[-l \ln(x_0 + \sqrt{x_0^2 - 1})]}{\sqrt{l}}.$$

Now it is not difficult to obtain the bounds on Φ_n. Letting $z = 1$ $(\theta = 0)$, we have

$$\Phi_n(s, 1) \leqslant \sum_l (2l+1)\, |a_l(s)|.$$

A principal contribution to the sum is provided by l_0 first partial waves, where l_0 is determined from the condition

$$s^N \exp[-l_0 \ln(x_0 + \sqrt{x_0^2 - 1})] \sim 1,$$

whence

$$l_0(s) \simeq \frac{N\sqrt{s}}{2m_\pi} \ln s.$$

Thus, we obtain

$$\frac{1}{\sigma_{ab}^n} \frac{d\sigma_{ab}^n}{d\cos\theta} \leqslant \frac{s}{4m_\pi^2} N^2 \ln^2 s.$$

For the angles $\theta \neq 0, \pi$, the estimate can be improved slightly once the inequality

$$|P_l(\cos\theta)| < \frac{2}{\sqrt{\pi(2l+1)\sin\theta}}$$

is used. In this case we have the following limitation:

$$\frac{1}{\sigma_{ab}^n(s)}\frac{d\sigma_{ab}^n}{d\cos\theta} \leqslant \frac{N^{3/2}}{3\sqrt{\pi\sin\theta}}\left(\frac{s}{m_\pi^2}\right)^{3/4}\ln^{3/2}s.$$

Whence the bounds for inclusive spectra are readily obtained:

$$\hat{f}_{ab\to cX}(s, \cos\theta)_{\theta=0,\pi} \leqslant \frac{s}{4m_\pi^2}N^2\langle n_c\rangle\sigma_{ab\to cX}\ln^2 s,$$

$$\hat{f}_{ab\to cX}(s, \cos\theta) \leqslant \frac{N^{3/2}}{3\sqrt{\pi\sin\theta}}\left(\frac{s}{m_\pi^2}\right)^{3/4}\langle n_c\rangle\sigma_{ab\to cX}\ln^2 s.$$

At small $x = 2k_\parallel/\sqrt{s}$ the scaling behavior of the spectra

$$f_{ab\to cX}(s, x, \mathbf{k}_\perp) \simeq f_{ab\to cX}^{as}(x, \mathbf{k}_\perp)$$

is known to be appreciably violated, since the inclusive spectra rise with s. Using the fact that $\theta = 0$, we have $\mathbf{k}_\perp = 0$ and $|\mathbf{k}| = k_\parallel$. It is possible therefore to rewrite the bounds for $f_{ab\to cX}$ in the following way:

$$\int_{-1}^{1}\frac{dx\, x^2}{\sqrt{x^2+\frac{4m_c^2}{s}}}f_{ab\to cX}(s, x, 0) \leqslant \text{const}\,\langle n_c\rangle\sigma_{ab\to cX}\ln^2 s.$$

This inequality depicts to what extent scaling can be violated without contradicting the theoretical generalities.

Using the unitarity condition, one can show that the functions

$$\hat{f}_{ab\to cX}^{(j)}(s, \cos\theta) = \int d|\mathbf{k}_c|\,\mathbf{k}_c^2 E_c^{j-1} f_{ab\to cX}(s, |\mathbf{k}_c|, \cos\theta)$$

continue analytically into the same domain where the function $\hat{f}_{ab\to cX}(s, \cos\theta) \equiv \hat{f}_{ab\to cX}^{(0)}(s, \cos\theta)$ is analytic as well. Consider the function

$$\hat{f}_{ab\to nC}^{(j)}(s, \cos\theta) = \int d|\mathbf{k}_c|\,\mathbf{k}_c^2 E_c^{j-1}\frac{d\sigma_{ab}^n}{d^3k_c}.$$

Let the coefficients $C_{mm'}^{ll'}(n, s, j)$ of its series expansion in Wigner functions be

$$\hat{f}_{ab\to nC}^{(j)}(s, \cos\theta) = \frac{2\pi}{\mathbf{k}_a^2}\sum_{l,l'}(2l+1)(2l'+1)$$
$$\times\sum_m C_{mm}^{ll'}(n, s, j)\, d_{m0}^l(\theta)\, d_{m0}^{l'}(\theta).$$

Then $C^{ll'}_{mm'}(n, s, j)$ can be related to the unitarity condition

$$\operatorname{Im} f_l(s) = |f_l(s)|^2 + \sum_m C^{ll}_{mm}(n, s, 0) + (\text{positive terms})$$

when it is multiplied by $(\sqrt{s})^j$. As a result, we obtain

$$(\sqrt{s})^j \operatorname{Im} f_l(s) \geqslant n_c \sum_m C^{ll}_{mm}(s, n, j),$$

where n_c is the number of the particles of sort c in the channel n.
Whence

$$|C^{ll'}_{mm}(n, s, j)| \leqslant \sqrt{C^{ll}_{mm}(n, s, j)\, C^{l'l'}_{mm}(n, s, j)} \leqslant \frac{(\sqrt{s})^j}{n_c} \times \sqrt{\operatorname{Im} f_l(s) \operatorname{Im} f_{l'}(s)},$$

which leads to the convergence of the expansion $\hat{f}^{(j)}_{ab\to nC}$ in the ellipse with foci at $\cos\theta = \pm 1$ and semi-major axis $x_0(s)$.

Whence, we can obtain bounds for the moments of the inclusive cross sections in terms of the energy of the detected particles at a fixed emission angle in the same manner as used to derive the bounds on $\hat{f}_{ab\to cX}$:

$$\hat{f}^{(j)}_{ab\to cX}(s, \cos\theta)|_{\theta=0,\pi} \leqslant \frac{s}{4m_\pi^2} N^2 (\sqrt{s})^j \sigma_{ab\to cX} \ln^2 s,$$

$$\hat{f}^{(j)}_{ab\to cX}(s, \cos\theta) \leqslant \frac{N^{3/2}}{3\sqrt{\pi\sin\theta}} \left(\frac{s}{m_\pi^2}\right)^{3/4} (\sqrt{s})^j \sigma_{ab\to cX} \ln^2 s,$$

where

$$\sigma_{ab\to cX}(s) = \sigma^{\text{tot}}_{ab}(s) \leqslant \frac{\pi}{m_\pi^2} \ln^2 s.$$

It follows, in particular, that

$$\int_{-1}^{1} dx\, |x|^{j+1} f_{ab\to cX}(s, x, 0) \leqslant \text{const}\cdot\ln^4 s.$$

The bounds of this type at $j = 1$ have been considered in [16]. However, for $\theta \neq 0, \pi$, the bounds given in this paper are singular too at $\theta = \pi/2$.

The bounds obtained at $j = 1$ can also be rewritten as a joint limitations on the mean density of the number of particles in transverse momentum, namely

$$\frac{d\langle n_c\rangle}{dk_\perp^2}(s, k_\perp^2) = \int dx \frac{d\langle n_c\rangle}{dx\, dk_\perp^2} = \frac{1}{\sigma_{ab\to cX}} \int \frac{dk_\parallel}{16\pi^2 E_c} f_{ab\to cX}(s, \mathbf{k}),$$

$$\int dk_\perp^2 \frac{d\langle n_c\rangle}{dk_\perp^2} = \langle n_c\rangle,$$

and on the mean value of $|x|^3$ for a fixed $k_\perp$

$$\langle |x|^3\rangle (s|k_\perp) = \frac{\int dx\, |x|^3\, d\langle n_c\rangle/dx\, dk_\perp^2}{\int dx\, d\langle n_c\rangle/dx\, dk_\perp^2},$$

when $k_\perp^2 = 0$

Indeed, when the limits at $\theta = 0$ and $\theta = \pi$ are summed and transformed to the variables x, $k_\perp^2$, we obtain

$$\int_{-1}^{1} dx\, x^2 x^* \frac{d\langle n_c\rangle}{dx\, dk_\perp^2} \leqslant \mathrm{const}\cdot \ln^2 s,$$

where $x^* = \sqrt{x^2 + 4m_c^2/s}$. Whence we have (since $x^* > |x|$)

$$\langle |x|^3\rangle (s|0) \frac{d\langle n_c\rangle}{dk_\perp^2}(s,\ 0) \leqslant \mathrm{const}\cdot \ln^2 s.$$

Since $\langle |x|^3\rangle \leqslant 1$, this is essentially a limit on the possible rise of $d\langle n_c\rangle/dk_\perp^2$ when $k_\perp^2 = 0$ and is dependent on the value of $\langle |x|^3\rangle (s|0)$.

For $j = 0$, we obtain a limit in terms of $\langle x^2\rangle (s|0)$ and $\langle n_c\rangle (s)$:

$$\left.\frac{d\langle n_c\rangle}{dk_\perp^2}\right|_{k_\perp^2=0} \leqslant \mathrm{const}\ \frac{\langle n_c\rangle \ln^2 s}{\langle x^2\rangle (s|0)}.$$

Thus, the angular properties in terms of the emission angle of the detected particle bring about rigorous inequalities which limit the mean density of the particle distribution of the transverse momentum $k_\perp$ as a function of the mean values of the quantities associated with longitudinal motion ($x = 2k_{\|}/\sqrt{s}$).

2.4.2 Upper Bounds Following from the Additional Hypotheses of Analyticity [17]

We have considered corollaries which follow from the analytical properties of the inclusive cross sections in terms of the emission angle of the detected particle. These, in turn, follow rigorously from the unitarity condition and analytic properties of the relevant amplitudes of elastic processes. Note that the position of the nearest singularity in the $\cos\theta$ plane ($x_0 \simeq 1 + 2m_\pi^2/s$) is indicative of the nearest momentum transfer singularity $t_{ac} = (k_a - k_c)^2$ when $t_{ac} = m_\pi^2$. This is a vivid demonstration of the short-range nature of nuclear forces between hadrons owing to an exchange of the lightest hadron which has the quantum numbers $a\bar{c}$ (for simplicity we shall assume this is a π-meson).

Just like about when we were dealing with the inclusive spectra integrated over the momentum modulus of the detected particle, x_0 (s)

corresponds to quasielastic processes. This is not difficult to see from the expression

$$t_{ac} = m_a^2 + m_c^2 - 2E_aE_c + 2\,|\mathbf{k}_a|\,|\mathbf{k}_c|\cos\theta.$$

The nearest real singularity in z stems from the maximum $|\mathbf{k}_c|$:

$$|\mathbf{k}_c| \simeq \sqrt{s}/2 \leqslant |\mathbf{k}_a| \simeq \sqrt{s}/2.$$

What is known of the analytical properties of the inclusive spectrum $f_{ab\to cX}(s, |\mathbf{k}_c|, \cos\theta)$ in $\cos\theta$ at fixed s and $|\mathbf{k}_c|$?

The only rigorously proved result is that there is analyticity within the Lehmann ellipse, the semi-major axis of which, $x_L(s)$, is independent of $|\mathbf{k}_c|$ as was established in Sec. 2.3.

It is natural to assume that like the case of the integrated spectrum, where the unitarity condition for the elastic process $a + b \to a + b$ enlarges the analyticity domain into an ellipse whose semi-major axis x_0 is greater than x_L, the unitarity condition for the process $a + b + c \to a + b + c$ will play a similar role for an inclusive spectrum at a fixed $|\mathbf{k}_c|$. However, even though a naive use of three-particle unitarity using phenomenological approaches (the Regge-Mueller analysis [5]) helps to provide some sort of unified description of various kinematic domains, rigorous analysis of this problem is as yet in embryo.

Nevertheless, relying on an analogy with the elastic processes, one can see that the decisive factor in this case will also be the short-range nature of nuclear forces associated with exchanging by a lightest hadron in the t-channel.

Without going into details on this account, assume, as it was in [18], that the nearest singularity in the z-plane is specified for $f_{ab\to cX}(s, |\mathbf{k}_c|, z)$ from the condition $t_{ac} = m_\pi^2$ as it was for the integrated spectrum. Whence it follows that the relevant analyticity domain as $s \to \infty$ and for fixed $|\mathbf{k}_c|$ is an ellipse with foci at $z = \pm 1$ and semi-major axis x_c which is approximately equal to the inverse velocity of the detected particle in C.M.S., i.e.

$$x_c \underset{s\to\infty}{\simeq} 1/v_c = \sqrt{1 + m_c^2/\mathbf{k}_c^2}.$$

What can be said about the singularities in the complex directions? Very little or nearly nothing, if one only considers the rigorous results. Yet we will, in addition to the foregoing, assume that there is an area in the vicinity of the physical domain of $\cos\theta$ where $f_{ab\to cX}(s, |\mathbf{k}_c|, \cos\theta)$ is analytical for all values of s and $|\mathbf{k}_c|$ which lie within the physical domain, viz. $s \geqslant (m_a + m_b)^2$, $0 \leqslant |\mathbf{k}_c| \leqslant \sqrt{s}/2$. To what extent does such an assumption fit the reality? When the analytic properties of the spectra, simulated both by individual diagrams [19] and by summed ladder amplitudes [20], were studied, it was found that the above hypothesis was actually realized in those cases.

Moreover, it proves that there are complex regions in the vicinity of the physical points of $\cos\theta$ which are free of the singularities at $s \to \infty$ and at any $|\mathbf{k}_c|$.

Thus, eventually, we assume that the inclusive spectrum $f_{ab\to cX}(s, |\mathbf{k}_c|, z)$ continues analytically in $z = \cos\theta$ into the domain $\mathscr{D}$ containing the ellipse of foci at $z = \pm 1$ and semi-major axis x_c, the sets $\mathscr{D} \cap \{z \mid \pm \operatorname{Im} z > 0\}$ being convex.

The condition of convexity, which means that the straight line segment joining any two points of the domain lies entirely inside the domain, is assumed here only to make the proof less unwieldy. In fact, the condition on the domain's form can be much less restrictive (see, e.g., [21]).

To produce the upper limits, it is more suitable to deal with the partially integrated spectrum

$$\bar{f}(s, p, z) = \int_0^p \frac{d|\mathbf{k}_c|\,\mathbf{k}_c^2}{8\pi^2 E_c} f_{ab\to cX}(s, |\mathbf{k}_c|\, z).$$

This is associated with the upper limits for $\bar{f}$, which contains less unknown information than does $f_{ab\to cX}(s, |\mathbf{k}_c|, z)$. At the same time we conserve the particular data on the form of the dependence of $f_{ab\to cX}$ on $|\mathbf{k}_c|$.

According to our fundamental assertions, the function $\bar{f}(s, p, z)$ is analytic in the domain $\mathscr{D}$, with the semi-major axis of the relevant ellipse being

$$\bar{x}_c = \sqrt{1 + m_c^2/p^2}.$$

Introduce a normalized function

$$F(s, p, z) = \frac{1}{a_0(s, p)} \bar{f}(s, p, z),$$

where

$$a_0(s, p) = \int_{|\mathbf{k}_c| \leqslant p} d^3k_c f_{ab\to cX}(s, \mathbf{k}_c).$$

Obviously, the analytic properties of F in z are the same as those of $\bar{f}$.

We expand $F(s, p, z)$ into a series in Legendre polynomials

$$F(s, p, z) = \sum_{l=0}^{\infty} (2l+1)\, b_l(s, p)\, P_l(z),$$

which converge uniformly and absolutely inside the ellipse with foci $z = \pm 1$ and semi-major axis $\bar{x}_c$. Note that

$$|b_l| = \frac{1}{2}\left|\frac{1}{a_0}\int dx\, P_l(z)\, \bar{f}(s, p, z)\right| < \frac{1}{2}.$$

To obtain the limitations on F, we use the technique of "associated functions" developed in [13, 21]. Let us introduce the following associated function:

$$G(s, p, \cos\theta, t) = \sum_{l=0}^{\infty} (2l+1)\, b_l(s, p)\, P_l(\cos\theta)\, t^l,$$

$$G(s, p, \cos\theta, 1) = F(s, p, \cos\theta).$$

As is shown in [13, 21], this function is analytic in t in the domain

$$\Delta_\theta = (\tilde{\mathscr{D}} e^{i\theta}) \cap (\tilde{\mathscr{D}} e^{-i\theta}),$$

where $\tilde{\mathscr{D}}$ is the domain containing a unit circle and confined by the boundary image of the domain $\mathscr{D}$ in the conformal mapping of the

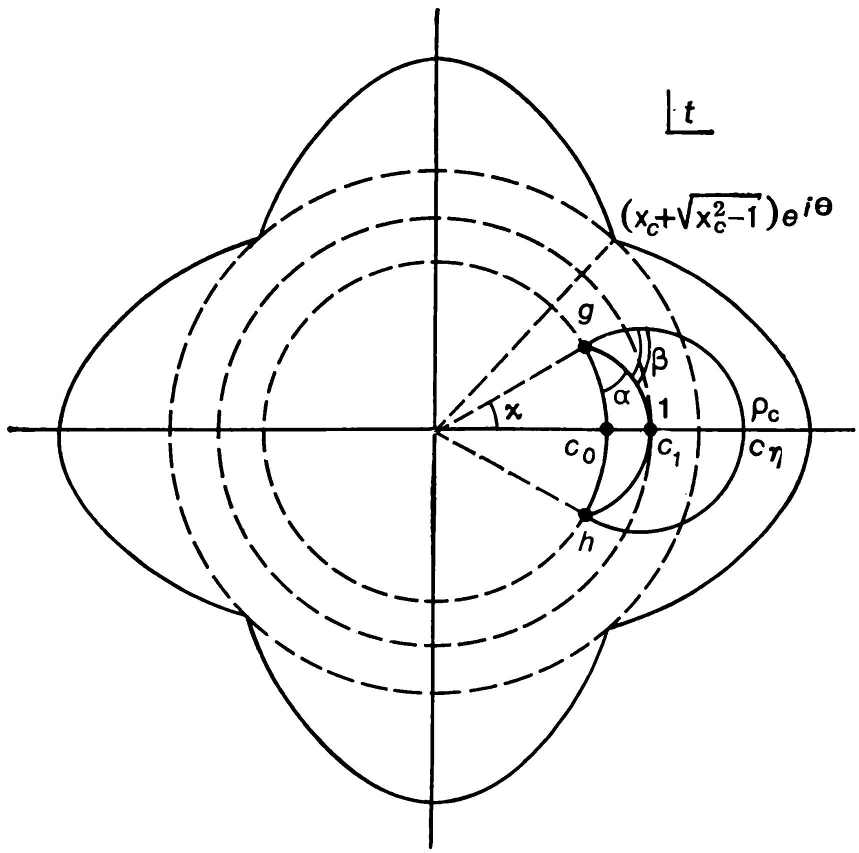


Fig. 2.1.

exterior of the segment (—1, 1) onto that of the unit circle (the inverse Zhukovskii transformation, see Fig. 2.1).

We draw a circle of radius $e^{-\zeta}$, $\zeta > 0$ in that plane and next, draw a circle through the point $\rho_c = x_c + \sqrt{x_c^2 - 1} + \eta \sin\theta$ whose

center lies on the real axis, which is orthogonal to the circle of radius $e^{-\zeta}$. At a sufficiently small η this circle lies in the domain Δ_θ.[4]

Denote the intersection points of these circles by g and h, and let C_0 and C_η be their arcs between g and h. Draw a circle through g, h, and $t = 1$. Let C_1 be the arc of this circle passing through $t = 1$.

Further let α be the angle between C_0 and C_1, β the angle between C_1 and C_η, and B the area confined by the arcs C_0 and C_η.

According to the two constants theorem, we have

$$F(s,\ p,\ \cos\theta) = G(s,\ p,\ \cos\theta;\ 1) \leqslant M_0^{\omega} M_\eta^{1-\omega}, \tag{2.21}$$

where

$$M_0 = \max_{t\in C_0} |G|,\quad M_\eta = \max_{t\in C_\eta} |G|,$$

and $\omega = \beta/(\alpha+\beta)$ is the magnitude of the harmonic measure $\omega\,(t, C_0, B)$ of the set C_0 relative to the domain B at the point $t = 1$. Assuming the polynomial boundedness of F in the domain $\mathscr{D}$, we have

$$|G|_{t\in\Delta_\theta} \leqslant \text{const}\cdot s^K,$$

where k is a positive constant. Whence it follows that

$$M_\eta \leqslant \text{const}\cdot s^K.$$

To estimate M_0, we can use an expansion of G in t, and require $|b_l| \leqslant 1/2$, from which we find that

$$M_0 \leqslant (1 - e^{-\zeta})^{-2}.$$

To obtain the best limit, we need to find a harmonic measure and to choose the appropriate value of ζ. It follows immediately from the construction that

$$\tan\alpha = \frac{2\tan\frac{\chi}{2}\tanh\frac{\zeta}{2}}{\tan^2\frac{\chi}{2} - \tanh^2\frac{\zeta}{2}},$$

$$\tan\frac{\chi}{2} = \frac{\rho_c - e^{-\zeta}}{\rho_c + e^{-\zeta}}.$$

Having chosen

$$\tanh\frac{\zeta}{2} = \tan\frac{\chi}{2}\,\frac{1}{\pi K \ln s}$$

---

[4] At the given point the assumption of convexity of the set $\mathscr{D} \cap \{z \mid \pm\operatorname{Im} z > 0\}$ is essential.

and taking into account that $\alpha + \beta = \pi/2$ by the construction, we find that the harmonic measure satisfies the inequality

$$1 - \omega = \frac{2\alpha}{\pi} < \frac{2}{\pi} \tan \alpha < \frac{1}{\pi K \ln s}.$$

Whence we also have that

$$1 - e^{-\zeta} = \frac{1}{K \ln s} \frac{\rho_c - 1}{\rho_c + 1}.$$

Substituting the values of the upper limits for M_0 and M_η into inequality (2.21), the above expressions taken into account, we obtain the following upper bounds for the inclusive spectrum:

$$\int_0^p \frac{d\,|\mathbf{k}_c|\,|\mathbf{k}_c|^2}{E_c} f_{ab\to cX}(s,\ |\mathbf{k}_c|,\ \cos\theta) \leqslant \text{const}\, a_0 \times \left[\frac{\rho_c + 1}{\rho_c - 1}\right]^2 \ln^2 s. \tag{2.22}$$

By analogy, one can show that

$$\int_0^p \frac{d\,|\mathbf{k}_c|\,\mathbf{k}_c^2}{E_c} E_c^j f_{ab\to cX}(s,\ |\mathbf{k}_c|,\ \cos\theta) \leqslant \text{const}\, a_j \left[\frac{\rho_c + 1}{\rho_c - 1}\right]^2 \ln^2 s,$$

where

$$a_j = \int_{|\mathbf{k}_c| \leqslant p} d^3 k_c E_c^j f_{ab\to cX}(s,\ \mathbf{k}_c).$$

These limits are valid and nontrivial at all values of θ. For $\theta \neq 0, \pi$ they can be slightly improved, once we make use of an inequality $|P_l(\cos\theta)| < 2/\sqrt{\pi(2l+1)\sin\theta}$ when deriving the estimate for M_0. In this case we have

$$\int_0^p \frac{d\,|\mathbf{k}_c|\,|\mathbf{k}_c|^2}{E_c} E_c^j f_{ab\to cX}(s,\ |\mathbf{k}_c|,\ \cos\theta) \leqslant \frac{\text{const}\, a_j \ln^{\frac{3}{2}} s}{\sqrt{\sin\theta}} \left[\frac{\rho_c + 1}{\rho_c - 1}\right]^{3/2}. \tag{2.23}$$

At the fixed p (the pionization region) we have in particular

$$\int_0^p d\,|\mathbf{k}_c|\,\mathbf{k}_c^2 f_{ab\to cX} \leqslant \text{const}\sqrt{s}\ln^4 s.$$

At $\theta \neq 0, \pi$ we obtain

$$\int_0^p d\,|\mathbf{k}_c|\,\mathbf{k}_c^2 f_{ab\to cX} \leqslant \text{const}\,\sqrt{s}\,\frac{\ln^{7/2} s}{\sqrt{\sin\theta}}.$$

Now let $p = \xi\sqrt{s}/2$. The integration can then only be performed over the fragmentation region, i.e. over the range of finite x and, as a consequence, we have

$$\int_{\xi_0}^{\xi} dx\, x^j f_{ab\to cX}(s,\ x,\ k_\perp = 0) \leqslant \text{const}\cdot \xi^j \ln^4 s.$$

Interestingly enough, the limits obtained allow for a significant (power-like) scaling violation in the pionization region while the admissible scaling violation in the fragmentation region is appreciably weaker (logarithmic). This can be considered as a qualitative agreement with experimental data.

Now we can write the limits in terms of the variables $k_\perp$ and rapidity $y = \frac{1}{2}\ln\frac{E+k_\parallel}{E-k_\parallel}$ which are often used in presenting the experimental data on inclusive cross sections. For $j = 1$, $\theta = \pi/2$, and $p = \sqrt{s}/2$, we obtain

$$\int dk_\perp k_\perp \left.\frac{d\langle n_c\rangle}{dy\,dk_\perp}\right|_{y=0} \leqslant \text{const}\sqrt{s}\ln^{3/2} s,$$

or

$$\langle k_\perp\rangle(s\mid y=0)\left.\frac{d\langle n_c\rangle}{dy}\right|_{y=0} \leqslant \text{const}\sqrt{s}\ln^{3/2} s,$$

where $\langle k_\perp\rangle(s\mid y)$ is the mean value of the transverse momentum of the detected particle at fixed y, and

$$\frac{d\langle n_c\rangle}{dy} = \int dk_\perp \frac{d\langle n_c\rangle}{dk_\perp dy} = \sum_n n_c \frac{1}{\sigma_{ab\to cX}}\frac{d\sigma_{ab}^n}{dy}$$

is the mean density of the number of detected particles in rapidity space.

Experiments demonstrate a noticeable rise of $d\langle n_c\rangle/dy$ for small y [22]. The mean transverse momenta also disclose a slow (in the accessible energy range) rise with energy.

The bound obtained above limits a potential rise in both the quantities $\langle k_\perp\rangle(s\mid 0)$ and $d\langle n_c\rangle/dy\,(s,\ 0)$ by $\sim\sqrt{s}\ln^{3/2}s$. Note that the "a priori" bounds have the form

$$\langle k_\perp\rangle(s\mid y) \leqslant \sqrt{s}/2,$$

$$\int dy\frac{d\langle n_c\rangle}{dy} = \langle n_c\rangle \leqslant \sqrt{s}/m_c.$$

For completeness, we shall give these bounds in terms of the variables E_c and $\eta = -\ln\tan(\theta/2)$ (the pseudorapidity), as they are also often used to present experimental data (mainly, on cosmic rays):

$$\langle E_c\rangle (s \mid \eta = 0) \left.\frac{d\langle n_c\rangle}{d\eta}\right|_{\eta=0} \leqslant \text{const}\cdot\sqrt{s}\ln^{3/2} s.$$

The "a priori" bounds are

$$\langle E_c\rangle (s \mid \eta) \leqslant \sqrt{s}/2,$$

$$\int d\eta \frac{d\langle n_c\rangle}{d\eta} = \langle n_c\rangle \leqslant \sqrt{s}/m_c.$$

The experimental quantity $d\langle n_c\rangle/d\eta$ at $\eta \sim 0$ rises rather fast with initial energy [22]. In much the same way as bound (4.3) was obtained, one could produce a bound for energy flow, i.e.

$$\frac{dE}{d\theta} = \sum_c \int \frac{d\,|\mathbf{k}_c|\,\mathbf{k}_c^2}{2(2\pi)^2} f_{ab\to cX}(s, |\mathbf{k}_c|, \theta) \frac{1}{\sigma_{ab}^{\text{tot}}} \sin\theta$$

where naturally,

$$\int_0^\pi d\theta \frac{dE}{d\theta} = E_a + E_b = E \quad (\text{in the C.M.S.,} \quad E = \sqrt{s}).$$

Then

$$\frac{1}{E}\frac{dE}{d\theta} \leqslant \text{const}\left[\frac{\sin\theta}{(2m_\pi/\sqrt{s} + \eta\sin\theta)^3}\right]^{1/2} \ln^{3/2} s.$$

The right-hand side of the inequality qualitatively reflects a situation observed experimentally, viz. the main portion of the initial energy flow passes into two jets collimated near the angles $\theta_+ \simeq \frac{m}{\sqrt{s}}$ and $\theta_- \simeq \pi - \frac{m}{\sqrt{s}}$ with a halfwidth of the order of θ_+.

Now let's go over to the range of high $p = \xi\sqrt{s}/2$ and θ. In this range (obviously one of high transverse momenta) we have

$$\int_0^{\bar{x}_\perp} dx_\perp x_\perp^j f_{ab\to cX}(s, x_\perp, \theta) \leqslant \text{const}\cdot\frac{\sin\theta\, x_\perp^{j-2}\sigma_{ab\to cX}}{s}\ln^{3/2} s$$

in terms of the variables θ, s, $x_\perp = 2k_\perp/\sqrt{s}$. Applying an extensively used parametrization of the inclusive cross sections in the high transverse momenta range, namely,

$$f_{ab\to cX} \simeq k_\perp^{-N}\varphi(x_\perp, \theta),$$

this inequality yields

$$N \geqslant 2.$$

The bound could, by and large, be saturated, yet only in the angular interval $\Delta\theta$, which degenerates to zero as the energy grows, i.e. in $[\theta_0 - \Delta\theta, \theta_0 + \Delta\theta]$, where $\theta_0 \neq 0, \pi$ and $\Delta\theta \sim (\ln s)^{-3/2}$. Note, however, that $\Delta\theta$ in our case degenerates substantially slower compared to a relevant interval that corresponds to bounds without the additional hypotheses of analyticity (Sec. 2.4.1) and for which $\Delta\theta \to 0$ exponentially. Can the obtained bounds be improved? Consider the function $f(s, p, z)$ as a simple example which meets all the above requirements:

$$f(s, p, z) = a_0(s, p)\frac{1}{4}\left(\frac{z_0+\sqrt{z_0^2-1}}{2}\right)^{1/2}[(z_0-z)^{-1/2} + (z_0+z)^{-1/2}],$$

where

$$z_0 = \frac{2E_aE_c+\mu^2-m_a^2-m_c^2}{2\,|\mathbf{k}_a|\,p} \to \bar{x}_c = \sqrt{1+m_c^2/p^2}.$$

The function $\bar{f}$ has the following properties:

1. $\int\limits_{-1}^{1} dz\,\bar{f} = a_0(s, p)$;

2. $\bar{f}$ is analytic in z over the entire complex plane having the cuts $(-\infty, -z_0]$, $[z_0, +\infty)$;

3. $\bar{f}$ is polynomially bounded in s at all p and z in the analyticity domain and so we have for large angles

$$\bar{f}(s, p, \cos\theta) \underset{s\to\infty}{\simeq} a_0(s, p)\frac{1}{4}\left(\frac{\bar{x}_c}{2}+\frac{m_c}{2p}\right)^{1/2} \times [(\bar{x}_c-\cos\theta)^{-1/2}+(\bar{x}_c+\cos\theta)^{-1/2}].$$

Thus, we see that using the assumptions made above we can improve the bounds only at the expense of lowering the exponent of $\ln s$.

2.4.3 The Lower Bound for the Decrease of Inclusive Spectrum with Increasing Transverse Momentum [17]

Above we established the upper bound to which the inclusive spectra could decrease at high transverse momenta. However, the true behavior of an inclusive spectrum can, certainly, differ strongly from the upper estimate. Here we shall establish the lower limit within the framework of the above assumptions about the analyticity in the emission angle of the detected particle.

By means in a conformal mapping

$$w(z) = x_c z^{-1}\left(x_c - \sqrt{x_c^2 - z^2}\right),$$

we can transform the z-plane with the cuts $(-\infty, -x_c]$, $[x_c, +\infty)$ into the interior of a circle of radius x_c. The domain $\mathscr{D}$ will be chosen so that by using the mapping $w(z)$ it is transformed into the w-plane inside the ellipse $\mathscr{E}$ (Fig. 2.2) which has foci at $\pm w(a)$ and a semi-major axis x_c. The parameter a can take any value from the interval

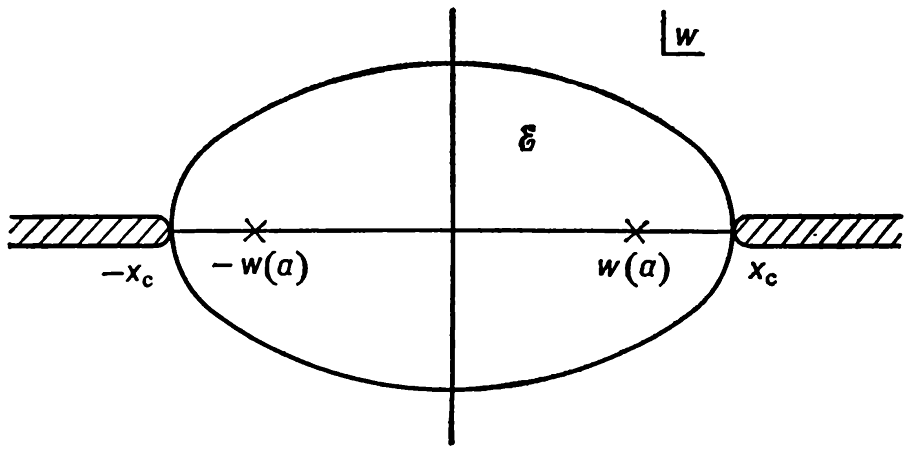


Fig. 2.2.

(0, 1). Note that the points $z = \pm 1$ are transformed into the points $\pm x_c$ $(x_c = \sqrt{x_c^2 - 1})$.

If we now use the transformation

$$\tau = \frac{1}{w(a)} [w + \sqrt{w^2 - w^2(a)}],$$

the domain $\mathscr{E}$ will be transformed into a ring of unit inner radius and outer radius R in the τ-plane,

$$R = \frac{1}{w(a)} [x_c + \sqrt{x_c^2 - w^2(a)}].$$

Note that a circle of unit radius in the τ-plane corresponds to the segment $[-w(a), w(a)]$ in the w-plane. An ellipse in the w-plane, with semi-major axis $w(1)$ and with foci $\pm w(a)$, when transformed in the τ-plane, becomes a circle of radius E,

$$E = \frac{1}{w(a)} [w(1) + \sqrt{w^2(1) - w^2(a)}].$$

The Hadamard theorem implies that

$$\max_{-a \leqslant \cos\theta \leqslant a} f_{ab \to cX}(s, |\mathbf{k}_c|, \cos\theta) \leqslant M_E \left(\frac{M_E}{M_R}\right)^{\frac{\ln E}{\ln (R/E)}},$$

where

$$M_E = \max_{z \in \Gamma(E)} |f_{ab \to cX}(s, |\mathbf{k}_c|, z)|,$$

$$M_R = \max_{z \in \Gamma(R)} |f_{ab \to cX}(s, |\mathbf{k}_c|, z)|.$$

Here $\Gamma(E)$ is the curve which passes through the points $z = \pm 1$ in the z-plane and whose image in the τ-plane is a circle of radius E, $\Gamma(R)$ being the boundary of the domain $\mathscr{D}$ in the z-plane. The points of the curve $\Gamma(R)$ which intercept the imaginary axis are

$$\operatorname{Im} z = \pm \frac{2\sqrt{2}\, a^2 x_c (x_c^2 - a^2)^{1/4}}{(x_c - \sqrt{x_c^2 - a^2})^{3/2}}.$$

By definition,

$$M_E \geqslant f_{ab\to cX}(s, |\mathbf{k}_c|, 1).$$

When

$$m_c \ll |\mathbf{k}_c| \leqslant \sqrt{s}/2,$$

it may be readily checked that

$$\frac{\ln E}{\ln R/E} = \frac{k_\perp^a}{m} \frac{\Lambda}{2\sin\theta} \ln \frac{1+\Lambda}{1-\Lambda} - 1,$$

where

$$\sin\theta_a = \sqrt{1-a^2}, \qquad \Lambda = \left(\frac{2\sin\theta_a}{1+\sin\theta_a}\right)^{1/2}, \qquad k_\perp^a = |\mathbf{k}_c| \sin\theta_a.$$

Note that

$$\varphi(\Lambda) \equiv \frac{\Lambda}{2\sin\theta_a} \ln \frac{1+\Lambda}{1-\Lambda} \geqslant 1.$$

The function $\varphi(\Lambda)$ grows slowly with θ_a. For example, in the angular domain ($0° \leqslant \theta_a \leqslant 80°$), φ changes from 1 to 3.

Thus, we have

$$\frac{\ln E}{\ln R/E} \simeq \frac{k_\perp^a}{2m_c} \frac{\Lambda}{\sin\theta_a} \ln \frac{1+\Lambda}{1-\Lambda} = \frac{k_\perp^a}{2m} \varphi(\theta_a).$$

The Hadamard inequality takes the form

$$\max_{|\cos\theta| \leqslant a} f_{ab\to cX}(s, |\mathbf{k}_c|, \cos\theta) \geqslant f_{ab\to cX}(s, |\mathbf{k}_c|, 1) \exp\left[\frac{-k_\perp^a}{m_c} \varphi(\theta_a) \ln \frac{M_R}{M_E}\right].$$

If an inclusive spectrum falls off monotonically with increasing θ, a maximum is reached at the boundary points $\cos\theta = \pm a$. Thus, for the case of identical initial particles (for example, the process $p + p \to \pi + X$) we have

$$f_{ab\to cX}(s, |\mathbf{k}_c|, \cos\theta) \geqslant f_{ab\to cX}(s, |\mathbf{k}_c|, 1) \exp\left[-\frac{k_\perp}{m_c} \varphi(\theta) \ln \frac{M_R}{M_E}\right].$$

Consider the range of angles θ such that

$$m_c \ll k_\perp \ll |\mathbf{k}_c| = x\sqrt{s}/2.$$

Then

$$f_{ab\to cX}(s,\ x,\ k_\perp) \geqslant f_{ab\to cX}(s,\ x,\ 0)\, e^{-B(s,\ x)\,k_\perp},$$

where $B(s,\ x) = \frac{1}{m_c}\ln\frac{M_R}{M_E}$. If $M_R \leqslant \text{const}\cdot s^N$, and

$$M_E \geqslant f_{ab\to cX}(s,\ x,\ 0) \geqslant \text{const}\cdot s^{-\nu},\quad \nu \geqslant 0,$$

then $B(s,\ x) \leqslant \text{const}\cdot\ln s$.

The lower limit obtained deviates badly from experimental data at very high ($\leqslant \sqrt{s}/2$) transverse momenta, where inclusive spectra level off.

Yet for reasonably high $k_\perp$ the lower boundary functionally coincides with the widely accepted parametrization exp $(-Bk_\perp)$. This fact seems especially valuable since the exponential fall off in $k_\perp$ at reasonably high $k_\perp$ is not entirely confirmed in theory, except for thermodynamic models.

2.5 THE 3 → 3 SCATTERING AMPLITUDE AND GENERALIZED OPTICAL THEOREM

The total cross section describing the process

$$a + b \to X$$

is simply related to a discontinuity in the forward scattering amplitude $(a + b \to a + b)$, viz.

$$\frac{1}{2i}[T_2(s+i0,\ 1) - T_2(s-i0,\ 1)] \equiv \operatorname{Im} T_2(s,\ 1)$$
$$= 2\sqrt{s}\,|\mathbf{k}_a|\,\sigma_{ab}^{\text{tot}}(s).$$

This ensues from the unitarity condition and the completeness of the system of physical states.

It is intuitively clear that the inclusive cross section of the process

$$a + b \to c + X \tag{2.24}$$

should likewise be related to some discontinuity in the forward scattering amplitude of the process

$$a + b + c \to a + b + c. \tag{I}$$

In [5] it is hypothesised that one of the discontinuities of the forward $3 \to 3$ scattering amplitude in the variable $M^2 = (k_a + k_b - k_c)^2$ coincides, to within the known factors, with an inclusive spectrum corresponding to process (I). This hypothesis received the name of the "generalized optical theorem". As will be seen below, the

discontinuity in the $3 \to 3$ forward amplitude in terms of a variable, which is linearly related to M^2, contains rather a large number of terms only some of which being identifiable with inclusive spectra [23]. Thus, in contrast to the $2 \to 2$ process, the $3 \to 3$ amplitude discontinuity has a more complicated form and is not expressible only in terms of some or other observable cross section.

2.5.1 A Modified Reduction Technique

The standard reduction technique [10] for process (I) yields retarded amplitudes of the type

$$\left\langle a \left| \frac{\delta^2 S^{(2)}(x_c, x_b)}{\delta\varphi_b^{\text{out}}(y_b)\,\delta\varphi_c^{\text{out}}(y_c)} \right| a \right\rangle,$$

where

$$S^{(2)}(x_c, x_b) = \frac{\delta^2 S}{\delta\varphi_c^{\text{out}}(x_c)\,\delta\varphi_b^{\text{out}}(x_b)} S^+.$$

Such an expression is insufficient in two respects:

(i) the variational derivative is nonzero in a domain which is too large for analytic continuation to be ensured for the variables we are interested in;

(ii) even granted that analytic continuation were possible, the absorptive amplitude part would be expressed via double current commutators and would not allow any simple physical interpretation. To overcome these difficulties, we shall exploit a technique based on applying both the variations in the out-fields, $\delta/\delta\varphi^{\text{out}}$, and the variations in the in-fields, $\delta/\delta\varphi^{\text{in}}$. The quantities related to the in-field will be shown with a tilde. The creation and annihilation operators are assumed to be

$$a_k^{\pm} \equiv a^{\pm}(\mathbf{p}_k),\quad \tilde{a}_k^{\pm} \equiv \tilde{a}^{\pm}(\mathbf{p}_k),\quad J_k \equiv J_k(x_k) = i\,\frac{\delta S}{\delta\varphi_k^{\text{out}}(x_k)} S^+.$$

$$\delta_k^{\pm} \equiv \delta/\delta\varphi_k^{\substack{\text{out}\\ \text{in}}}(x_k),$$

$$\Gamma_k^{\pm} F(x_k) = \int dx_k e^{\pm i p_k x_k} F(x_k) \Big|_{p_k^0 = \sqrt{\mathbf{p}_k^2 + m_k^2}}.$$

With these notations, a matrix element of the $m \to n$ transition is written in the form

$$\langle n \mid S \mid m \rangle = \langle n \mid \widetilde{m} \rangle - \langle n \mid \tilde{a}_1^+ \mid \widetilde{m-1} \rangle.$$

The operators $\tilde{a}_1^+$ and a_1^+ are linked by the standard relationship

$$\tilde{a}_1^+ = S a_1^+ S^+ = a_1^+ - i\Gamma_1^- J_1.$$

Whence

$$\langle n \mid S \mid m\rangle = \langle n \mid a_1^+ \mid m\rangle - i\Gamma_1^- \langle n \mid J_1 \mid \widetilde{m-1}\rangle.$$

As a result, the physical matrix element $\langle n \mid S \mid m\rangle$ is expressed through the matrix element of the current J_1 between the states in the different bases. The term $\langle n \mid a_1^+ \mid \widetilde{m-1}\rangle$ stands for the sum of the disconnected contributions to the amplitude.

Subsequent reduction steps consist in successive commutation of the current J_1 (or its variation) with the "increation" operators and "out-annihilation" operators according to the rules

$$[a_k^-, F] = \Gamma_k^+ \delta_k^+ F,$$

$$[F, \tilde{a}_k^+] = \Gamma_k^- \delta_k^- F,$$

where

$$\delta_k^- F[\varphi^{\text{out}}] = S[\delta_k^+(S^+ F S)] S^+ = \delta_k^+ F - i[F, J_k]. \tag{2.25}$$

From the consistency condition

$$\delta_k^+ J_l - \delta_l^+ J_k = i[J_l, J_k]$$

it follows that the operations $\delta_k^+ \delta_l^+$ (δ_k^-, δ_l^-) are commutative, and (2.25) yields the relation

$$\delta_k^- J_l = \delta_l^+ J_k. \tag{2.26}$$

Therefore, the causality condition in terms of δ^- is written as

$$\delta_k^- J_l = 0, \quad x_k \gtrsim x_l,$$

i.e. "oppositely" with respect to the causality condition in terms of δ^+:

$$\delta_k^+ J_l = 0, \quad x_l \gtrsim x_k.$$

Thus, apart from the occurrence of trivial "single-particle" terms, the main outcome of the reduction proves to be matrix elements of the "generalized retarded operators"

$$\delta_n^{\pm} \delta_{n-1}^{\pm} \ldots \delta_2^{\pm} J_1.$$

In the above,

(i) a domain, where the generalized retarded operators vanish due to the causality condition, is substantially larger than the appropriate domains of major radiation operators, and this essentially improves the possibility of analytic continuation;

(ii) the absorptive part of the $m \to n$ amplitude is expressed in terms of simple current commutators and their variations and allows a more vivid interpretation than would be possible with the ordinary reduction method.

2.5.2 Construction of Retarded Amplitude

To proceed further, we must choose the independent variables and three are presented here. We will study the analytic properties of one of them, while retaining the two other variables to be real and fixed within the definite intervals of their admissible values. As a frame of reference, we will select the laboratory system with

$$\mathbf{p}_a = 0.$$

When the 4-vectors Δ and Q are introduced,

$$\Delta = \frac{1}{2}(p_c - \alpha p_b),$$

$$Q = \frac{1}{2}(1-\eta)\, p_c + \frac{1}{2}(1+\eta)\, p_b,$$

where

$$\alpha = p_c^0 / p_b^0, \quad \eta = \frac{m_c^2 - m_b^2 \alpha}{-4m^2}, \quad m^2 = -\Delta^2,$$

it is evident that

$$\Delta_0 = 0, \ \ (Q\Delta) = -(\mathbf{Q}\cdot\mathbf{\Delta}) = 0.$$

When $Q_0 = \omega$, α, m^2 are selected as independent variables, it can be seen Q^2 is independent of ω:

$$Q^2 = m_c^2 + m^2(1+\eta) = \alpha^2 m_b^2 + m^2(1-\eta)^2.$$

In terms of these variables the momenta p_c and p_b have the following form:

$$p_c = (\omega, |\mathbf{Q}|\,\mathbf{e}_1 + (1+\eta)\, m\mathbf{e}_2),$$

$$p_b = \frac{1}{\alpha}(\omega, \ |\mathbf{Q}|\,\mathbf{e}_2 - (1-\eta)\, m\mathbf{e}_2),$$

where

$$|\mathbf{Q}| = \sqrt{\omega^2 - Q^2} = \sqrt{\omega^2 - m_c^2 - m^2(1+\eta)^2}, \quad \mathbf{e}_1 = \frac{\mathbf{Q}}{|\mathbf{Q}|},$$

$$\mathbf{e}_2 = \frac{\mathbf{\Delta}}{|\mathbf{\Delta}|}, \quad \mathbf{e}_1 \cdot \mathbf{e}_2 = 0.$$

The physical domain of the variables ω, α, m is prescribed by the inequalities

$$\omega \geqslant \sqrt{Q^2} = \sqrt{m_c^2 + m^2(1+\eta)^2} \equiv \omega_{\mathrm{thr}}$$

$$0 < \alpha < \infty,$$

$$0 < m^2 < \infty.$$

The amplitude of process (I) is continued analytically in the variable ω at the fixed α and m.

To exclude disconnected contributions of the type $\langle n \mid a_1^+ \mid \overparen{m-1}\rangle$, we shall assume the reduction takes place when $\mathbf{p}_j \neq \mathbf{p}_j'$ for any matching pairs of initial and final momenta, and only in the final expressions put $\mathbf{p}_j = \mathbf{p}_j'$. Then all disconnected contributions proportional to $\delta(\mathbf{p}_j - \mathbf{p}_j')$ will vanish.

Following the technique described in Sec. 2.5, we give the starting matrix element of process (I) in the form

$$\langle a'b'c' \mid S \mid abc\rangle = \langle a'b'c' \mid \tilde{a}_b^+ \mid \overparen{ac}\rangle$$
$$= -i\Gamma_b^- \langle a'b'c' \mid J_b \mid \overparen{ac}\rangle + \langle a'b'c' \mid a_b^+ \mid \overparen{ac}\rangle.$$

The second term vanishes by virtue of the preceding assertion. Further reduction yields

$$\Gamma_b^- \langle a'b'c' \mid J_b \mid \overparen{ac}\rangle = \Gamma_b^- \Gamma_{b'}^+ \langle a'c' \mid \delta_b^+ J_{b'} \mid \overparen{ac}\rangle$$
$$+ \Gamma_{b'}^- \langle a'c' \mid J_b a_{b'}^- \mid \overparen{ac}\rangle,$$
$$\Gamma_b^- \Gamma_{b'}^+ \langle a'c' \mid \delta_{b'}^+ J_b \mid \overparen{ac}\rangle$$
$$= \Gamma_b^- \Gamma_{b'}^+ \Gamma_c^- \langle a'c' \mid \delta_c^- \delta_{b'}^+ J_b \mid a\rangle + \Gamma_b^- \Gamma_{b'}^+ \langle a'c' \mid \tilde{a}_c^+ \delta_{b'}^+ I_b \mid a\rangle.$$

We took into account here the stability of the single particle state:

$$\mid \tilde{a}\rangle = S \mid a\rangle = \mid a\rangle.$$

Finally, the result of the reduction looks like

$$i\langle a'b'c' \mid S \mid abc\rangle = \Gamma_b^- \langle a'c' \mid J_b a_{b'}^- \mid \overparen{ac}\rangle$$
$$+ \Gamma_b^- \Gamma_{b'}^+ \langle a'c' \mid \tilde{a}_c^+ \delta_{b'}^+ J_b \mid a\rangle + \Gamma_b^- \Gamma_{b'}^+ \Gamma_c^- \Gamma_{c'}^+ \langle a' \mid \delta_{c'}^+ \delta_c^- \delta_{b'}^+ J_b \mid a\rangle \tag{2.27}$$

Excluding the unrelated part, the first term in (2.27) can be presented as follows:

$$\Gamma_b^- \langle a'c' \mid J_b a_b^- \mid \overparen{ac}\rangle = -i\Gamma_b^- \Gamma_{b'}^+ \langle a'c' \mid J_b J_{b'} \mid \overparen{ac}\rangle.$$

Substituting a complete system of intermediate states with translational invariance taken into account, we obtain

$$-i \sum_N (2\pi)^4 \delta(p_{a'} + p_{c'} - p_b - p_N)\, \delta(p_a + p_c - p_{b'} - p_N)$$
$$\times \langle a'c' \mid J_b(0) \mid N\rangle \langle N \mid J_{b'}(0) \mid \widetilde{ac}\rangle.$$

This expression does not vanish only when

$$(p_a + p_c - p_{b'})^2 \geqslant M^2, \quad p_a^0 + p_c^0 - p_{b'}^0 \geqslant 0,$$

where M^2 is the least mass of the intermediate states contributing to $\sum\limits_N \ldots$.

Substituting a parametrization of p_b and p_c in terms of ω, α, and m^2 in this condition, one can readily see that $\sum\limits_N \ldots = 0$ at

$$\omega < \omega_{\max} = \frac{\alpha (m_c^2 - M^2)}{2m_a (1-\alpha)} - \frac{1}{2m_a} (m_c^2 - \alpha m_b^2) - \frac{2m^2}{m_a (1-\alpha)} .$$

Whence it follows that the first term in (2.27) vanishes in the physical domain $\omega > \omega_{\text{thr}}$, given that $\omega_{\max} < \omega_{\text{thr}}$. It is evident that this will be deliberately so, provided $\omega_{\max} < 0$, which, in turn, complies with the condition

$$m^2 > \frac{\alpha}{4} (m_a^2 - M^2) - \frac{1}{4} (1-\alpha)(m_c^2 - \alpha m_b^2). \tag{2.28}$$

In the case of $m_b = m_c = 0$ (for example, b and c are photons), we obtain

$$\begin{aligned} &m_a^2 = M^2, \\ &\omega_{\text{thr}} = m, \\ &\omega_{\max} = -\frac{2m^2}{m_a (1-\alpha)}, \\ &\omega_{\max} < \omega_{\text{thr}}, \end{aligned} \tag{2.29}$$

invariably at $m > 0$, $\alpha < 1$.

Now consider the second term in (2.27). Dropping the disconnected contributions, as always, and using the stability of the single particle states, we have

$$\begin{aligned} &\Gamma_b^- \Gamma_{b'}^+ \langle a'c' \mid \tilde{a}_c^+ \delta_{b'}^+ J_b \mid a \rangle = -i \Gamma_b^- \Gamma_{b'}^+ \Gamma_c^- \langle a'c' \mid J_c \delta_{b'}^+ J_b \mid a \rangle \\ &= -i \Gamma_b^- \Gamma_{b'}^+ \Gamma_c^- \Gamma_{c'}^+ \langle a' \mid \delta_{c'}^+ J_c \delta_{b'}^+ J_b \mid a \rangle \\ &-i \Gamma_c^- \Gamma_{c'}^+ \Gamma_b^- \Gamma_{b'}^+ \langle a' \mid J_c \delta_{c'}^+ \delta_{b'}^+ J_b \mid a \rangle . \end{aligned} \tag{2.30}$$

When the term which contains $\langle a' \mid J_c \delta_{c'}^+ \delta_{b'}^+ J_b \mid a \rangle$ is expanded in complete system of the intermediate states, it yields $\delta (p_{a'} - p_c - k_N)$ for every intermediate state, and hence it vanishes, since it cannot, for any N, meet the condition $p_{a'} - p_c - k_N = 0$ by virtue of particle stability.

An analogous procedure results in the identity

$$\begin{aligned} &\Gamma_c^- \Gamma_{c'}^+ \Gamma_b^- \Gamma_{b'}^+ \langle a' \mid \delta_{c'}^+ J_c \delta_{b'}^+ J_b \mid a \rangle \\ &= \int d^3k \langle a'c' \mid S^+ \mid c\mathbf{k} \rangle \langle b'\mathbf{k} \mid S \mid ab \rangle + \sum_{N \geq 2} \ldots , \end{aligned}$$

where the terms with multiparticle intermediate states contain $\delta (p_{a'} + p_{c'} - p_c - k_N)$. For forward scattering when $p_c = p_{c'}$,

only a single-particle intermediate state with $\mathbf{k} = \mathbf{p}_{a'} = \mathbf{p}_a$ is possible, and the remaining terms vanish.

Yet the contribution of the single-particle intermediate state describes the process of twofold two-particle rescattering $(a + b \to$ $\to "k" + b) \otimes ("k" + c \to a + c)$ and contains a singularity of the form $\delta\,(k_a^2 - m_a^2)$. To study the analytical properties of true three-particle scattering amplitude, such contributions must be specially isolated.

Eventually, when the terms in the forward scattering vanish in a part of the physical domain that is bounded by conditions (2.29), the connected part of the starting matrix elements reads:

$$\langle a'c'b' \mid S \mid acb\rangle = -i\Gamma_c^- \Gamma_{c'}^+ \Gamma_b^- \Gamma_{b'}^+ \langle a' \mid \Delta^{\mathrm{ret}}(x_c, x_b, x_{c'}, x_{b'}) \mid a\rangle,$$

where

$$\Delta^{\mathrm{ret}}(x_c, x_b, x_{c'}, x_{b'}) = r_{\mathrm{I}}^{\mathrm{ret}}(x_c, x_b, x_{c'}, x_{b'}) - i\delta_c^+ J_{c'} \delta_{b'}^+ J_b,$$

$$r_{\mathrm{I}}^{\mathrm{ret}}(x_c, x_b, x_{c'}, x_{b'}) = \delta_{c'}^+ \delta_c^- \delta_{b'}^+ J_b.$$

Introduce a connected retarded amplitude of process (I):

$$\langle a'b'c' \mid S \mid abc\rangle = i\,(2\pi)^4\,\delta\,(p_{a'} + p_{b'} + p_{c'} - p_a - p_b - p_c)$$
$$\times\, T_{\mathrm{I}}^{\mathrm{ret}}(p_{a'} p_{b'} p_{c'};\, p_a p_b p_c).$$

Here $T_{\mathrm{I}}^{\mathrm{ret}}$ is related to Δ^{ret} in the following way:

$$T_{\mathrm{I}}^{\mathrm{ret}}(p_{a'}, p_{b'}, p_{c'};\, p_a, p_b, p_c)$$
$$= \int dy_c\, dy_{c'}\, dy_{b'}\, e^{-i(p_c y_c - p_{c'} y_{c'} - p_{b'} y_{b'})} \langle a' \mid \Delta^{\mathrm{ret}}(y_c\, 0\, y_{c'} y_{b'}) \mid a\rangle,$$

where

$$y_c = x_c - x_b, \quad y_{c'} = x_{c'} - x_b, \quad y_{b'} = x_{b'} - x_b.$$

For forward scattering,

$$T_{\mathrm{I}}^{\mathrm{ret}}(p_a p_b p_c)$$
$$= \int dy_c\, dy_{c'}\, dy_{b'}\, e^{ip_c(y_{c'} - y_c) + ip_b y_{b'}} \langle a \mid \Delta^{\mathrm{ret}}(y_c\, 0\, y_{c'} y_{b'}) \mid a\rangle. \tag{2.31}$$

As mentioned above, to study the analytic properties in the variable ω we must isolate a term related to $\langle a' \mid \delta_c^- J_{c'} \delta_b^- J_{b'} \mid a\rangle$ which describes the rescattering from $T_{\mathrm{I}}^{\mathrm{ret}}$. The remaining part will be denoted as $R_{\mathrm{I}}^{\mathrm{ret}}$ and called the reduced retarded amplitude. It can be represented as

$$R_{\mathrm{I}}^{\mathrm{ret}} = \int dy_c\, dy_{c'}\, dy_{b'}\, e^{ip_c(y_{c'} - y_c) + ip_b y_{b'}} \langle a \mid r_{\mathrm{I}}^{\mathrm{ret}}(y_c\, 0\, y_{c'} y_{b'}) \mid a\rangle. \tag{2.32}$$

2.5.3 Causality Condition

In the laboratory frame ($\mathbf{p}_a = 0$) the exponent in (2.32) takes the form

$$i\,(\omega z^0 - |\mathbf{Q}|\,(\mathbf{e}_1\cdot\mathbf{z})) \\ -im\left[(1+\eta)\,(\mathbf{e}_2\cdot(\mathbf{y}_c - \mathbf{y}_{c'}) - \frac{(1-\eta)}{\alpha}\,(\mathbf{e}_2\cdot\mathbf{y}_{b'}))\right],$$

where

$$z = y_{c'} - y_c + \frac{1}{\alpha}\,y_{b'}.$$

Let us find domain, where the operator $r_{\mathrm{I}}^{\mathrm{ret}}\,(y_c \circ y_{c'}\,y_{b'})$ does not vanish. Above all, due to relation (2.26), $r_{\mathrm{I}}^{\mathrm{ret}}$ can be rewritten as

$$r_{\mathrm{I}}^{\mathrm{ret}}\,(y_c \circ y_{c'} y_{b'}) = \delta_{c'}^{+}\delta_{c}^{-}\delta_{b}^{-}J_{b'}.$$

The commutativity of δ_c^- and δ_b^- and the causality condition suggest that $r_{\mathrm{I}}^{\mathrm{ret}} \neq 0$ except when

$$y_{b'} > y_c,\quad y_{b'} > 0, \tag{2.33}$$

where $x > y$ signifies $(x - y)^2 > 0$ and $x^0 > y^0$.

In order to find the limits associated with the variable y_c, let us transform $r_{\mathrm{I}}^{\mathrm{ret}}$ with the help of (2.25) to

$$\delta_{c'}^{+}\delta_{c}^{-}\delta_{b'}^{+}J_b - \delta_{c}^{-}\delta_{c'}^{+}\delta_{b'}^{+}J_b = -i\,[\delta_{b'}^{+}J_b,\ \delta_{c'}^{+}J_c] \\ = -[\delta_{b'}^{+},\ \delta_{b}^{-}]\,\delta_{c'}^{+}J_c = -\delta_{b'}^{+}\delta_{b}^{-}\delta_{c'}^{+}J_c + \delta_{b}^{-}\delta_{b'}^{+}\delta_{c'}^{+}J_c.$$

The causality condition leads to $r_{\mathrm{I}}^{\mathrm{ret}} \neq 0$ when

$$y_{c'} > 0 \text{ or } y_{c'} > y_c. \tag{2.34}$$

Whence it follows that $r_{\mathrm{I}}^{\mathrm{ret}} \neq 0$ in the union of the domains Σ_1 and Σ_2:

$$\Sigma_1 : y_{b'} > y_c,\quad y_{b'} > 0,\quad y_{c'} > y_c,$$
$$\Sigma_2 : y_{b'} > y_c,\quad y_{b'} > 0,\quad y_{c'} > 0.$$

In each of the domains Σ_1 and Σ_2, the 4-vector z can be presented as the sum of components, each of which belongs to a closed future cone, $\overline{V}_+ = \{x \mid x_0 > |\mathbf{x}|\}$. In the domain Σ_1

$$z = (y_{c'} - y_c) + \frac{1}{\alpha}\,y_{b'} \in \overline{V}_+ \text{ at } \alpha > 0,$$

and in domain Σ_2

$$z = y_{c'} + (y_{b'} + y_c) + \frac{1-\alpha}{\alpha} y_{b'} \in \bar{V}_+,$$

when $0 < \alpha < 1$.

Thus, everywhere inside the support of $r_{\mathrm{I}}^{\mathrm{ret}}$ the vector z lies inside the future cone, $\bar{V}_+$, if only $0 < \alpha < 1$.

2.5.4 Advanced Amplitude

An advanced amplitude results from a reduction of the matrix element $\langle a'b'c' \mid S^+ \mid abc \rangle$. Repeating the foregoing procedure almost word-in-word, we arrive at an expression like (2.31), where the δ^+ are changed into δ^- and vice versa, and in place of Δ^{ret} there is an operator,

$$\Delta^{\mathrm{adv}} = \delta_{c'}^- \delta_c^+ \delta_{b'}^- J_b + i\delta_{c'}^- J_c \delta_{b'}^- J_b,$$

and

$$T_{\mathrm{I}}^{\mathrm{adv}} = \int dy_c\, dy_{c'}\, dy_{b'} e^{ip_c(y_c - y_{c'}) + ip_b y_{b'}}$$
$$\times \langle a \mid \Delta^{\mathrm{adv}} (y_c\, 0\, y_{c'} y_{b'}) \mid a \rangle.$$

The expression for Δ^{adv} can be reduced to the form

$$\Delta^{\mathrm{adv}} (y_c\, 0\, y_{c'} y_{b'}) = r_{\mathrm{I}}^{\mathrm{adv}} (y_c\, 0\, y_{c'} y_{b'}) + i\delta_{b'}^- J_b \delta_{c'}^- J_c,$$

where

$$r_{\mathrm{I}}^{\mathrm{adv}} = \delta_c^+ \delta_{c'}^- \delta_{b'}^- J_b,$$

whence the reduced advanced amplitude is equal to

$$R_{\mathrm{I}}^{\mathrm{adv}} (p_a p_b p_c)$$
$$= \int dy_c\, dy_{c'}\, dy_{b'} e^{ip_c(y_{c'} - y_c) + ip_b p_{b'}} \langle a \mid r_{\mathrm{I}}^{\mathrm{adv}} (y_c\, 0\, y_{c'} y_{b'}) \mid a \rangle.$$

Since the change of $\delta^+ \to \delta^-$ leads to the substitution of the future cones, $\bar{V}_+$, for the past comes, $\bar{V}_- = \{x \mid x_0 < - \mid x \mid\}$, then like the foregoing it is possible to see that everywhere in the domain, where $r_{\mathrm{I}}^{\mathrm{adv}} \neq 0$, the vector $z = y_{c'} - y_c - \frac{1}{\alpha} y_{b'}$ lies inside $\bar{V}_-$ given that $0 < \alpha < 1$.

2.5.5 An Absorptive Part of the Reduced Amplitudes

The absorptive part of the reduced amplitude of process (I),

$$A_{\mathrm{I}} = \frac{1}{2i} [R_{\mathrm{I}}^{\mathrm{ret}} - R_{\mathrm{I}}^{\mathrm{adv}}],$$

can be expressed as a sum of the matrix elements of the current commutators and their variational derivatives.

Having transformed from the variant δ^+ to those of δ^- in the operators $r_{\mathrm{I}}^{\mathrm{ret,adv}}$, the following expression can easily be obtained:

$$A_{\mathrm{I}} = -\sum_{i=1}^{6} A_{\mathrm{I}i}, \tag{2.35}$$

where

$$\begin{aligned}
&A_{\mathrm{I}i} = \langle a' \mid [\mathcal{O}_1^i, \mathcal{O}_2^i] \mid a\rangle,\\
&\mathcal{O}_1^1 = \delta_{c'}^- J_b, \quad \mathcal{O}_2^1 = \delta_c^- J_{b'},\\
&\mathcal{O}_1^2 = \delta_c^- J_b, \quad \mathcal{O}_2^2 = \delta_{c'}^- J_{b'},\\
&\mathcal{O}_1^3 = J_b, \quad \mathcal{O}_2^3 = \delta_c^- \delta_{c'}^- J_{b'},\\
&\mathcal{O}_1^4 = \delta_c^- \delta_{c'}^- J_b, \quad \mathcal{O}_2^4 = J_{b'},\\
&\mathcal{O}_1^5 = \delta_b^- \delta_c^- J_{b'}, \quad \mathcal{O}_2^5 = J_{c'},\\
&\mathcal{O}_1^6 = J_c, \quad \mathcal{O}_2^6 = \delta_{b'}^- \delta_{c'}^- J_b.
\end{aligned}$$

Thus, while the absorptive part of the reduced amplitude for the $2 \to 2$ process contains the contributions from two channels, for the $3 \to 3$ process it contains the contributions from twelve channels.

Nevertheless, even in this, more complicated, case there is a "gap", i.e. an interval of the values of ω, in which the absorptive part vanishes.

In fact, $A_{\mathrm{I}i}$ can be presented in the form

$$\begin{aligned}
A_{\mathrm{I}i} = &\sum_N (2\pi)^4 \delta(k_N - P_1^i) \int dy_c\, dy_{c'} e^{ip_c(y_{c'} - y_c)}\\
&\times \langle a \mid \mathcal{O}_1^i \mid N\rangle \langle N \mid \mathcal{O}_2^i \mid a\rangle\\
&- \sum_N (2\pi)^4 \delta(k_N - P_2^i)\\
&\times \int dy_c\, dy_{c'} e^{ip_c(y_{c'} - y_c)} \langle a \mid \mathcal{O}_2^i \mid N\rangle \langle N \mid \mathcal{O}_1^i \mid a\rangle,
\end{aligned}$$

where

$$\begin{aligned}
&P_1^1 = p_a + p_c - p_b, \quad P_2^1 = p_a + p_b - p_c,\\
&P_1^2 = p_a - p_b - p_c, \quad P_2^2 = p_a + p_b + p_c,\\
&P_1^{3,4} = p_a - p_b, \quad P_2^{3,4} = p_a + p_b,\\
&P_1^{5,6} = p_a - p_c, \quad P_2^{5,6} = p_a + p_c.
\end{aligned}$$

If a is a nucleon, and b, c are mesons, then $A_{1i} \neq 0$ at

$$-\infty < \omega < -\omega_i, \quad \omega_i < \omega < \infty, \quad \omega = \pm\omega_{iP},$$

where

$$\omega_1 = (2m_a(1-\alpha))^{-1}[\alpha((m_a+\mu)^2 - m_a^2 - m_b^2 - m_c^2) + (4m^2 + m_c^2 + \alpha^2 m_b^2)],$$

$$\omega_2 = (2m_a(1-\alpha))^{-1}[\alpha((m_a+\mu)^2 - m_a^2 - m_b^2 - m_c^2) - (4m^2 + m_c^2 + \alpha^2 m_b^2)],$$

$$\omega_{3,4} = \alpha(2m_a)^{-1}[(m_a+\mu)^2 - m_a^2 - m_b^2],$$

$$\omega_{5,6} = (2m_a)^{-1}[(m_a+\mu)^2 - m_a^2 - m_c^2],$$

μ is the π-meson's mass, and ω_{iP} corresponds to the poles and differs from ω_i by the substitution of $m_a + \mu \to m_a$.

It is not difficult to see that when $0 < \alpha < 1$ and $m^2 > 0$, the leftmost of the right thresholds is $\omega_2 : \omega_i \geqslant \omega_2$. Therefore, the condition of gap availability reduces to the condition that $\omega_2 > 0$, i.e.

$$0 < 4m^2 < \alpha(2m_a\mu + \mu^2) - (1 + \alpha)(m_c^2 + \alpha m_b^2).$$

Gap availability, however, does not as yet assure the derivation of dispersion relations. It is in addition essential that $R_{\mathrm{I}}^{\mathrm{ret}}$ and $R_{\mathrm{I}}^{\mathrm{adv}}$ continue into the appropriate half-planes of the complex quantity ω. Further, the positions of the thresholds ω_i in the general case can go beyond the limits of the physical domain, i.e. $\omega_i < \omega_{\mathrm{thr}}$. Then amplitude discontinuities in the domain $\omega_i < \omega < \omega_{\mathrm{thr}}$ have no physical sense.

There is, however, a situation when the nonphysical domain is absent. Let the particles b and c have zero mass, e.g. b and c are photons, and the process

$$\gamma + \gamma + a \to \gamma + \gamma + a \tag{II}$$

is treated at the lowest ($\sim e^4$) order. In this special case

$$\omega_{\mathrm{thr}} = m,$$

and the nonphysical domain is absent when

$$0 < 4m < \sqrt{(1+\alpha)^2 m_a^2 + 4\alpha(2m_a\mu + \mu^2)} - (1+\alpha)m_a,$$

with the condition simultaneously ensuring gap availability as well.

As is proved in [23], the dispersion relation occurs here (for simplicity, we will drop the polar terms and do not explicitly indicate the possible subtractions):

$$\frac{1}{2}[R_{\mathrm{II}}^{\mathrm{ret}}(\omega, \alpha, m) + R_{\mathrm{II}}^{\mathrm{adv}}(\omega, \alpha, m)] = \frac{1}{\pi}\int_{\omega_2(\alpha, m)}^{\infty} d\omega' \frac{2\omega' A_{\mathrm{II}}(\omega', \alpha, m)}{\omega'^2 - \omega^2}.$$

2.5.6 Generalized Optical Theorem

Consider that the discontinuity A_{III} complies with the second term in (2.35). Substituting $\delta^+_{b'}J_c$ and $\delta^+_b J_{c'}$ for $\delta^-_c J_{b'}$ and $\delta^-_{c'} J_b$ in it, we arrive at the inclusive spectrum of the process $\gamma + a \to \gamma' + X$, viz.

$$f_{\gamma a\to\gamma X}(\omega,\ \alpha,\ m)=\frac{1}{I}A_{\mathrm{II}\,1}(\omega,\ \alpha,\ m)$$

$$=\frac{1}{I}\sum_N (2\pi)^4\,\delta\,(k_N-p_a-p_\gamma+p_{\gamma'})$$

$$\times\left|\int dy_{\gamma'}e^{ip_{\gamma'}y_{\gamma'}}\left\langle N\left|\frac{\delta J_\gamma(y_{\gamma'})}{\delta\varphi_\gamma(0)}\right|a\right\rangle\right|^2,$$

where $I = m_a\omega/\alpha$ is the flux of initial particles. So, of all the terms on the right-hand side of (2.35), only one, namely A_{III} conforms to the physical characteristics (to the inclusive process cross section). The remaining terms have no immediate interpretation via observed quantities.

In order for this result to be exploited in practice it is obviously necessary to "single-out" those terms in the unitarity condition for the $3 \to 3$ process which are associated with the inclusive processes, i.e. to make the surplus terms vanish. The procedure to do such a singling-out has not as yet been developed, and it will perhaps be influenced by the analytic properties of the remaining variables (α, m).

In practice (the study of the asymptotic behavior of inclusive spectra, for example), one actually deals with a term of the complete discontinuity that conforms to the inclusive process and applies, say, the Regge formulas for asymptotics in some kinematic range to it

2.6 THE ASYMPTOTIC BEHAVIOR OF INCLUSIVE CROSS SECTIONS AND THE UNIVERSAL NATURE OF STRUCTURE FUNCTIONS

As has been shown in Secs. 2.3 and 2.4, the requirements of microcausality and spectrality are important in establishing the analytic properties of multiple production amplitudes in terms of angular variables and the resultant limitations on the behavior of the inclusive cross sections at high energies.

On the other hand, when the asymptotic properties of the inclusive cross sections were studied a number of regularities were revealed in the behavior of the inclusive spectra under specific asymptotic regimes. These are, for example, the approximate scaling of structure functions in deep inelastic lepton-nucleon scattering, limiting

fragmentation, or Feynman scaling in the pure hadron inclusive processes.

To what extent do these regularities agree with the fundamentals of local quantum field theory and, in particular, with the principles of microcausality and spectrality?

This item has in part been tackled in the preceding sections, where the bounds on the possible scaling violations in hadron inclusive processes were obtained.

In this section we shall proceed in a more straightforward way, i.e. directly from the relevant Jost-Lehmann-Dyson (JLD) representations for the matrix elements of the (retarded) commutators of the local current operators, without resorting to the analytic properties of the amplitudes. The amplitudes, cross sections, or structure functions, written in terms of the JLD representation satisfy the microcausality and spectrality principles provided some assumptions are made not for the amplitudes, cross sections, or structure functions, but rather for the appropriate spectral functions, to be sure, without violating their general properties thereby. Acting in this way, we can, on the one hand, describe the asymptotic physical characteristics possible, that at least do not contradict theory, and on the other check for the agreement of these fundamentals with various hypotheses with respect to these asymptotic properties of the inclusive processes.

2.6.1 Deep Inelastic Lepton-Hadron Scattering

To be definitive (yet without any loss of generality), consider the deep inelastic process of an electron being scattered by a nucleon:

$$e + N \to e + X.$$

In the single-photon exchange approximation the cross section of this process is defined using the Fourier transform of the matrix element of the commutator $j_\mu(x)$ of electromagnetic currents:

$$W_{\mu\nu} = \frac{1}{8\pi} \int d^4x\, e^{iqx} \sum_\sigma \langle \mathbf{p},\ \sigma |\, [j_\mu(x),\ j_\nu(0)]\, | \mathbf{p},\ \sigma \rangle,$$

where q is the 4-momentum of the virtual photon ($q^2 < 0$), and ($\mathbf{p}$, σ) the nucleon's momentum and spin. The deep inelastic scattering cross section can be expressed in terms of the structure functions $W_{1,2}$ which are related to the $W_{\mu\nu}$ in the following fashion:

$$W_{\mu\nu} = \left(-g_{\mu\nu} + \frac{q_\mu q_\nu}{q^2} \right) W_1$$
$$+ \left(p_\mu - \frac{qp}{q^2}\, q_\mu \right) \left(p_\nu - \frac{qp}{q^2}\, q_\nu \right) W_2.$$

The microcausality principle suggests that

$$\widetilde{W}_{\mu\nu}(x, p) = \sum_{\sigma} \langle \mathbf{p}, \sigma | [j_\mu(x), j_\nu(0)] | \mathbf{p}, \sigma \rangle = 0$$

when $x^2 < 0$. On the other hand, by inserting the full system of intermediate states between the current operators into the expression for $W_{\mu\nu}$, we obtain that

$$W_{\mu\nu}(q, p) = 0 \quad \text{for} \quad p^0 \pm q^0 < \sqrt{m^2 + (\mathbf{p} \pm \mathbf{q})^2},$$

where m is the nucleon's mass. This corresponds to the spectrum of the intermediate states that yield non-vanishing contributions to $W_{\mu\nu}$ beginning from the single-nucleon state.

According to Jost, Lehmann and Dyson [24], the functions having these properties of disappearing in coordinate and momentum spaces, are parametrized in the following way (in the $\mathbf{p} = 0$ frame):

$$W_i(q, p) = \int d\mathbf{u}\, d\lambda^2 E(q')\, \delta(q_0 - (\mathbf{q} - \mathbf{u})^2 - \lambda^2)\, \psi_i(|\mathbf{u}|, \lambda^2),$$

where ψ_i''s vanish outside the domain

$$\{\mathbf{u}, \lambda^2 |\ |\mathbf{u}| \leqslant m, \ \lambda^2 \geqslant (m - \sqrt{m^2 - \mathbf{u}^2})\}.$$

The JLD representation has been used in [25] to analyze rigorously the compatibility of the automodel behavior of the structure functions of deep inelastic scattering to the fundamentals of quantum field theory, i.e. microcausality and spectrality.

Without going into further details of the analysis performed in [25], we shall only exemplify its idea briefly by a concrete specific case. It is easily seen that, once ψ_2 is intergratable over λ^2, then as $q^2 \to -\infty$ and when $x = -q^2/(2qp) = -q^2/\nu$ and kept fixed,

$$\nu W_2(x, q^2) = \int d\mathbf{u}\, \delta(x - \mathbf{e} \cdot \mathbf{u}) \int d\lambda^2 \psi_2(|\mathbf{u}|, \lambda^2) = F_2(x),$$

where $\mathbf{e} = \mathbf{q}/|\mathbf{q}|$, i.e. at the Bjorken limit, $\nu W_2(x, q^2)$ tends to the scale invariant limit $F_2(x)$. Likewise, if as $\lambda^2 \to \infty$,

$$\psi_1(|\mathbf{u}|, \lambda^2) \to \psi_1^{(0)}(|\mathbf{u}|),$$

then $W_1(x, q^2) \to F_1(x)$. In the general case the spectral functions, $\psi(\mathbf{u}, \lambda^2)$, are generalized functions of a moderate growth, i.e. they grow moderately fast as $\lambda^2 \to \infty$ and/or $|\mathbf{u}| \to 0$. For example, as $\lambda^2 \to \infty$, the inequality

$$\int d\mathbf{u} \varphi(\mathbf{u})\, \psi(\mathbf{u}, \lambda^2) \leqslant \text{const}\, \lambda^{2k}, \quad k > 0$$

should be satisfied (the integral $\int d\mathbf{u}\varphi\psi$ having the test function φ being treated using the theory of generalized functions).

Thus, the causality and spectrality principles expressed by means of the JLD representation do not contradict rigorous Bjorken scaling.

Bjorken scaling is known to be realized only approximately as the structure functions are q^2-dependent in the various ranges of the variable x values in a different fashion (albeit relatively weakly). The compatibility of these, more general, asymptotics and theory has rigorously been analyzed on the basis of the theory of generalized functions in [26].

2.6.2 Hadron-Hadron Inclusive Processes

Consider the amplitude of the $a+b\rightarrow nC$ multiple process contributing to the nth channel of the inclusive process

$$a+b\rightarrow c+X.$$

Using the Bogolyubov reduction formula, this amplitude can be presented in the form

$$\langle n|T|ab\rangle = \frac{-1}{\sqrt{n_c}}\int dx\, e^{i\frac{k_a+k_b}{2}x}\left\langle n'\left|\frac{\delta J_a\left(-\frac{x}{2}\right)}{\delta\varphi_c^+\left(\frac{x}{2}\right)}\right|b\right\rangle.$$

According to the microcausality condition,

$$\delta J_a\left(-\frac{x}{2}\right)\Big/\delta\varphi_c^+\left(\frac{x}{2}\right)=0,$$

when $x_0<|\mathbf{x}|$. Whence it follows that

$$\langle n|T|ab\rangle=\frac{1}{2\pi\sqrt{n_c}}\int_{-\infty}^{+\infty}dQ_0\,\frac{F_n\left(Q_0,\ \frac{\mathbf{k}_a+\mathbf{k}_c}{2}\right)}{Q_0-\frac{1}{2}(k_a^0+k_c^0)-i0},$$

where

$$F_n(Q_0,\mathbf{Q})=\int d^4x\, e^{iQx}\left\langle n'\left|\left[J_a\left(-\frac{x}{2}\right),\ J_c\left(\frac{x}{2}\right)\right]\right|b\right\rangle$$
$$=\sum_m \delta(Q-\Delta_a-P_m)\langle n'|J_a|m\rangle\langle m|J_c|b\rangle$$
$$-\sum_{m'}\delta(Q+\Delta_a-P_{m'})\langle n'|J_c|m'\rangle\langle m'|J_c|b\rangle,$$
$$n'=n-(0,\ \ldots,\ 1_c,\ \ldots,\ 0);\quad P_m=\sum_{il}k_{lc_i};\quad 1\leqslant l\leqslant m_i.$$

It is easy to see that in the frame of reference $\Delta_a = \mathbf{k}_b + \frac{\mathbf{k}_a - \mathbf{k}_c}{2} = 0$ $F_n(Q) = 0$ in the domain

$$\Delta_a^0 - \sqrt{\mathbf{Q}^2 + m_1^2} \leqslant Q_0 \leqslant \sqrt{\mathbf{Q}^2 + m_2^2} - \Delta_{a0}, \tag{2.36}$$

where $2\Delta_{a0} = 2k_b^0 + k_a^0 - k_c^0$, m_1 is the mass of the lowest state $|m\rangle$, for which

$$\langle n' | J_a | m\rangle \langle m | J_c | b\rangle \neq 0,$$

and m_2 is the mass of the lowest state $|m'\rangle$, for which

$$\langle n' | J_c | m'\rangle \langle m' | J_a | b\rangle \neq 0.$$

So. $F_n(Q)$ vanishes in domain (2.36), as does its Fourier image $\widetilde{F}_n(x)$ outside the light cone, since

$$\left[J_a\left(-\frac{x}{2}\right), J_c\left(\frac{x}{2}\right)\right] = 0$$

when $x^2 < 0$.

The following JLD representation is valid for this sort of domain

$$F_n(Q) = \int d^4v\, d\varkappa^2 \varepsilon(Q^0 - v^0)\, \delta[(Q-v)^2 - \varkappa^2]\, \Phi_n(v, \varkappa^2).$$

Moreover, the spectral function Φ_n vanishes outside the domain

$$|\mathbf{v}| \leqslant \Delta_a^0, \quad |\mathbf{v}| - \Delta_a^0 \leqslant v^0 < \Delta_a^0 - |\mathbf{v}|,$$

$$\varkappa \geqslant \max\{0, m_1 - \sqrt{(v^0 - \Delta_a^0)^2 - \mathbf{v}^2}, m_2 - \sqrt{(v^0 + \Delta_a^0)^2 - \mathbf{v}^2}\}.$$

If $m_1 = m_2 = m$, the integration over v may be chosen on the plane $v^0 = 0$, for which the representation for $F_n(Q)$ has the form

$$F_n(Q) = \varepsilon(Q^0) \int d\mathbf{v}\, d\varkappa^2 \delta(Q_0^2 - (\mathbf{Q} - \mathbf{v})^2 - \varkappa^2)[\Phi_{1n} + Q^0 \Phi_{2n}], \tag{2.37}$$

and the spectral functions $\Phi_{n1,2}$ vanish outside the domain $\Omega(\mathbf{v}, \varkappa^2)$:

$$|\mathbf{v}| \leqslant \Delta_a^0, \quad \varkappa \geqslant \max\{0, m - \sqrt{\Delta_a^{02} - \mathbf{v}^2}\}.$$

Note that the definition of F_n can be extended appropriately to make the support of F_n symmetric in Q_0, and so we shall from now on use representation (2.37). Thus the following representation is obtained for the amplitude $\langle n | T | ab\rangle$:

$$\langle n | T | ab\rangle = \int d\mathbf{v}\, d\varkappa^2 \frac{\Phi_{n1}(\mathbf{v}, \varkappa^2) + Q^0 \Phi_{n2}(\mathbf{v}, \varkappa^2)}{(Q_0 + i\varepsilon)^2 - (\mathbf{Q} - \mathbf{v})^2 - \varkappa^2},$$

where

$$Q = \frac{1}{2}(k_a + k_c).$$

Whence the JLD representation for the inclusive spectrum of the process $a + b \to c + X$ can be obtained as

$$f_{ab\to cX} = I^{-1}(s)\int d\mathbf{v}_1 d\varkappa_1^2\, d\mathbf{v}_2\, d\varkappa_2^2$$
$$\times [(Q^0+i\varepsilon)^2-(\mathbf{Q}-\mathbf{v}_1)^2-\varkappa_1^2]^{-1}[Q^0-i\varepsilon)^2-(\mathbf{Q}-\mathbf{v}_2)^2-\varkappa_2^2]^{-1}$$
$$\times [\Phi_1(\mathbf{v}_1,\varkappa_1^2,\mathbf{v}_2,\varkappa_2^2;\ k_b,\ k_a-k_c)+Q^0\Phi_2+Q^{02}\Phi_3],$$

where

$$I(s)=\sqrt{(k_ak_b)^2-m_a^2m_b^2},$$
$$\Phi_1+Q^0\Phi_2+Q^{02}\Phi_3=\sum_n (2\pi)^4\,\delta(k_a+k_b-k_c-P_{n'})$$
$$\times[\Phi_{n1}(\mathbf{v}_1\varkappa_1^2)+Q^0\Phi_{n2}(\mathbf{v}_1\varkappa_1^2)]^*[\Phi_{n1}'(\mathbf{v}_2\varkappa_2^2)+Q^0\Phi_{n2}(\mathbf{v}_2\varkappa_2^2)].$$

Φ_i obviously vanishes outside the domain

$$\Omega(\mathbf{v}_1\varkappa_1^2)\times\Omega(\mathbf{v}_2\varkappa_2^2).$$

As the independent variables which characterize the inclusive process, we choose

$$s=(k_a+k_b)^2,\ M^2=(k_a+k_b-k_c)^2,\ t=(k_a-k_c)^2.$$

If the momenta Q, k_b, and $k_a - k_c$ in the chosen frame of reference $(2\mathbf{k}_b+\mathbf{k}_a-\mathbf{k}_c=0)$ are expressed in terms of these spectral invariants, the spectral functions will prove to be independent of s and the only s-dependence (via Q) will be contained in the representation's denominators and in the explicity isolated factors Q^0 and Q^{02}.

Consider therefore, an asymptotic domain

$$s \gg M^2,\ t.$$

In the laboratory system $(\mathbf{k}_b = 0)$ it conforms to fast particles, viz.

$$E_c=E_a-\frac{M^2-t-m_b'}{2m_b},\quad E_a\simeq\frac{s}{2m_b},$$

moving in the direction of the incident beam m inside the angular interval

$$\theta \leqslant \frac{\sqrt{-t}}{E_a}.$$

In the C.M.S. we have, respectively,

$$E_c^*=\frac{\sqrt{s}}{2}-\frac{M^2-m^2}{2\sqrt{s}},$$
$$\theta^*\simeq 2\sqrt{\frac{-t}{s}}\ll 1,$$

i.e. qualitatively the same situation; viz. the fast particles moving almost in the direction of the incident particle. Thus, the domain

chosen by us is contained in the fragmentation region of the particle a.

Assume now that the integrals ($i = 1, 2, 3$)

$$V_i = \int \frac{d\mathbf{v}_1 \, d\varkappa_1^2 \, d\mathbf{v}_2 \, d\varkappa_2^2}{[(\mathbf{v}_1 \cdot \mathbf{e}) - i0] \, [(\mathbf{v}_2 \cdot \mathbf{e}) + i0]} \Phi_i (\mathbf{v}_1, \varkappa_1^2; \mathbf{v}_2, \varkappa_2^2; M^2, t)$$

converge ($\mathbf{e} = \lim_{s \to \infty} \mathbf{Q} / | \mathbf{Q} |$). The asymptotics of the inclusive spectrum will now have the form

$$f_{ab \to cX} (s, t, M^2) \simeq \frac{1}{4s} V_3 (t, M^2).$$

The above convergence assumptions are too restrictive and do not usually hold true (e.g., in perturbation theory). The spectral functions can grow with $\varkappa_i^2$ (though certainly not faster than some power of $\varkappa_i^2$, this being implied by the fact that the Φ_i's are generalized functions of moderate growth). Besides, singularities can be encountered on the surface $\mathbf{v}_i \mathbf{e} = 0$; for example, singularities like $| \mathbf{v}_i \mathbf{e} |^{-\alpha}$ (with the case of $\alpha = 0$ being singular, too). All integrals, it must therefore be kept in mind, are understood in the sense of the theory of generalized functions, i.e. subtractions are done that make all integrals converge.

Suppose the spectral function Φ_3 can be exemplified in the vicinity of the surface $(\mathbf{v}_i \cdot \mathbf{e}) = 0$ as

$$\Phi_3 \simeq \prod_{i=1}^{2} (\mathbf{v}_i \cdot \mathbf{e})^{-\alpha} \Phi_3 (\varkappa_i^2, \ldots), \quad \alpha > 0,$$

where $\Phi_{1,2}$ are everywhere regular and

$$\int \frac{d\varkappa_1^2 \, d\varkappa_2^2}{\varkappa_1^2 \varkappa_2^2} \Phi_3 (\varkappa_1^1, \varkappa_2^2, \ldots) <^{?} \infty,$$

then the asymptotics of the inclusive spectrum will be given by the following expression:

$$f_{ab \to cX} (s, t, M^2) \cong \left(\frac{s}{s_0} \right)^{2\alpha(t, M^2) - 2} R_c (t, M^2).$$

Here we have explicitly allowed for the fact that the index α could be generally dependent on t and M^2 and that s_0 can be chosen in an arbitrary fashion and be dependent on t, M^2 as well. The explicit form of $R_c (M^2, t)$ is specified by the function Φ_3.

The asymptotic form of $f_{ab \to cX} (s, t, M^2)$ can also be influenced by the possible growth in $\varkappa_i^2$, e.g. $\sim (\varkappa_i^2)^\beta (ln \, \varkappa_i^2)^\sigma$, where β and σ are also, generally speaking, dependent on t and M^2.

If no dynamic speculations are added relative to the evolution of the spectral functions, then the theory alone tells us nothing about

the singularities which are essential to the spectral functions. The only guide in studying this problem is that spectral functions are the generalized functions of moderate growth, i.e. they cannot grow in $\varkappa_i^2$ faster than a polynomial in $\varkappa_i^2$, and the singularities cannot be too strong in $\mathbf{v}_i$. For example, behavior like $(\mathbf{v}_i\cdot\mathbf{e})_+^{-\alpha}\ln^{\beta}(\mathbf{v}_i\cdot\mathbf{e})$ is admissible, but a singularity like $\exp|\mathbf{v}_i\cdot\mathbf{e}|^{-1}$ is not. The powers of the singularities must not be such as will violate the limits to the inclusive spectra that follow from the sum rules, analyticity, and unitarity which were treated above.

A detailed analysis of a wide class of spectral functions' singularities that are admissible in the above sense has been presented in [27]. As a result, a general form was obtained for the asymptotes of the inclusive processes in the fragmentation region of the particle a, which is consistent with the causality and spectrality principles:

$$f_{ab\to cX}(s,\ t,\ M^2)$$
$$\simeq \sum_j \left(\frac{s}{s_0}\right)^{\eta_j(t,\ M^2)}\left[\ln\left(\frac{s}{s_0}\right)\right]^{\sigma_j(t,\ M^2)} F_j(t,\ M^2)$$
$$+\text{possible interference terms}.$$

The sum rules

$$\sum_c \int d^3k_c E_c f_{ab\to cX} = 2(2\pi)^3(E_a+E_b)\,\sigma_{ab}^{\text{tot}}$$

and the Froissart bound

$$\sigma_{ab}^{\text{tot}} \leqslant \text{const}\ \ln^2\left(\frac{E_a+E_b}{E_0}\right)$$

lead to

$$\max_j \eta_j(t,\ M^2)\leqslant 1,$$
$$\sigma_j(t,\ M^2)\leqslant 2+n_j,$$

where the n_j's are specified by the behavior of $F_j\ (t,\ M^2)$ in the vicinity of the points t_0, M^2, and where

$$\max_{t,\ M^2} \eta_j(t,\ M^2)=\eta_j(t_0,\ M_0^2).$$

In the particle's fragmentation region in ($s \gg u = (k_b - k_c)^2$, M^2), i.e. where the detected particle has a high energy (in the $\mathbf{p}_a = 0$ frame), viz.

$$E_c \simeq E_b - (M^2 - u - m_a^2)/(2m_a),$$

and moves approximately in the direction of the particle b, similar conclusions can be drawn, i.e.

$$f_{ab\to cX}(s, u, M^2)$$
$$\simeq \sum_j \left(\frac{s}{s_0}\right)^{\bar{\eta}_j(u, M^2)} \left[\ln\left(\frac{s}{s_0}\right)\right]^{\bar{\sigma}_j(u, M^2)} \bar{F}_j(u, M^2)$$
$$+ \text{possible interference terms}$$

with the limitations on $\bar{\eta}_j(u, M^2)$, $\bar{\sigma}_j(u, M^2)$ being the same as with the fragmentation region of the particle a.

The indices η_j, σ_j, $\bar{\eta}_j$, $\bar{\sigma}_j$ are governed by the indices of the singularities of the spectral functions in the variables ν_i and $\varkappa_i^2$.

In the particular case of the binary process

$$a + b \to c + d,$$

$f_{ab\to cX} \sim \delta(M^2 - m_d^2)$ and, by integrating over M^2, we obtain similar asymptotic expressions for the differential cross sections of the forward $(d\sigma/dt)$ and backward $(d\sigma/du)$ cones.

Note that the asymptotic formulas produced contain the Regge asymptotics as a special case, provided η_i is independent of M^2 and corresponds to the Regge trajectories $\alpha(t)$ $(\alpha(u))$. Here the functions F_j and $\bar{F}_j$ are related to the residues of the relevant Regge trajectories and for the inclusive process to a triple-reggeon vertex (at $M^2 \gg \gg m^2_{a,b,c}$) as well.

2.6.3 The Universal Nature of Structure Functions in Inclusive Hadron-Hadron and Deep Inelastic Lepton-Hadron Processes

We have considered above cross sections for the lepton-hadron deep inelastic and hadron-hadron inclusive processes using a unified method based on the JLD representation. In this way, the cross sections were, in addition to the s-dependence, also described by the structure functions W_i, for deep inelastic inclusive processes, and by the F_j, for hadron-hadron inclusive processes. It has been hypothesized [7] that in both these cases the structure functions related to the same target could have analogous properties and within this framework a unified method of describing processes of different origin can be made in terms of the appropriate structure functions. Let us elucidate this point in more detail.

Deep inelastic processes, for example

$$e + p \to e + X,$$

are described by the structure functions $W_i(\nu, q^2)$, which are dependent on the transferred energy and momentum of leptons.

The structure functions W_i do not depend on the overall momentum of leptons, since the lepton-photon (W-boson) interaction Lagrangian is local, and the interaction in the electromagnetic (weak) coupling constant is taken into account only in the lowest order approximation of perturbation theory.

When hadrons interact in the inclusive process

$$a + b \to c + X,$$

even though the initial strong interaction Lagrangian is local, we cannot as yet neglect the nontrivial dependence on $k_a + k_c$ because of higher approximations in the coupling constant, so that the locality disappears and the inclusive spectrum now becomes essentially dependent on $k_a + k_c$, rather than just on $k_a - k_c$.

Thus, because of the hadron's internal structure, the effective interaction of the hadrons a and c is nonlocal and the inclusive spectrum is essentially dependent on three variables, e.g. $\nu = 2k_b\,(k_a - - k_c)$, $q^2 = (k_a - k_c)^2$, and $k_b\,(k_a + k_c)$. Therefore, the descriptions of the inclusive lepton-hadron and hadron-hadron processes seem, generally speaking, essentially different.

On the other hand, the kinematics of the deep inelastic processes is entirely applicable to the domain considered here, $s \gg M^2 = q^2 + \nu$, $t = -q^2$, given ν, $|q^2| \gg m^2$.

In the case of lepton-hadron deep inelastic scattering ($a = e$, $c = e$, $b = p$), the operators

$$\delta J_e(-x/2)/\delta\bar{\psi}_e(x/2)$$
$$= \delta\left\{T\left[\bar{\psi}_e\left(-\frac{x}{2}\right)A_\mu^{(0)}\left(-\frac{x}{2}\right)\gamma^\mu S\right]S^+\right\}\Big/\delta\bar{\psi}_e'\left(\frac{x}{2}\right)$$

are approximated to a high degree of accuracy by the following term in the series expansion of perturbation theory (for lepton fields):

$$\delta J_e(-x/2)/\delta\psi_e(x/2) \simeq A_\mu^{(0)}\gamma^\mu\delta(x),$$

where $A_\mu(x) = T\,[A_\mu^{(0)}(x)\,S]\,S^+$ is the Heisenberg operator for the electromagnetic field.

An inclusive spectrum for any process can be given in the form

$$f_{ab\to cX} = \frac{1}{s}\int dx\, e^{iqx}\langle p\,|I(x;\;Q)\,I^+(0;\;Q)|\,p\rangle,$$

where

$$I(x;\;Q) = \int dz\, e^{iQz}\,\delta J_a(x - z/2)/\delta\varphi_c^+(x + z/2),$$

$$p = k_b,\quad Q = \frac{1}{2}(k_a + k_c).$$

The operator $I(x; Q)$ is generally nonlocal in x. In the process $e + p \to e + X$ the operator $I(x; Q)$ is "localized", and we arrive at the inclusive spectrum

$$f_{ep\to eX} \simeq \sum_{i=1}^{2} \left(\frac{s}{s_0}\right)^{\eta_j} W_i(\nu, q^2),$$

where W_1 and W_2 are ordinary structure functions and the indices η_j are independent both of ν and q^2. We see that this is a special case of the general formula obtained in Sec. 2.6.2, where the η_j are, generally speaking, dependent on q^2 and ν. A concrete form of the factor governing the s-dependence is apparently influenced by the interaction details in a vertex binding a and c. If the interaction is local, the power of s is independent of ν and q^2, and if not, such a dependence comes into being. It is only natural to think that the cross sections for deep inelastic processes and pure hadronic processes essentially differ only in the s-factor.

It is possible to put forward an important physical assumption that hadron interactions are described in an inclusive process by the same type of structure functions that describe deep inelastic processes [7]. Thus structure functions are universal and can characterize processes of different nature. This does not exclude the structure functions for deep inelastic and pure hadronic inclusive processes from possessing a number of properties in common.

Nowadays this viewpoint is widespread and agrees both with experimental data and the results, for example, of the parton model.

According to the above hypothesis, the effective locality of an interaction should formally comply with the generalized structure functions, i.e.

$$W_{\alpha\beta}(\nu, q^2) = \int dx\, e^{iqx} \langle p \mid J_\alpha(x) J_\beta^\dagger(0) \mid p \rangle,$$

where $J_\alpha(x)$ are the local "currents" carrying the quantum numbers $\bar{a}c$ (isospin, hypercharge, etc.). In a special case the currents $J_\alpha(x)$ can coincide with electromagnetic or weak currents.

Thus, not only can we "probe "the electromagnetic or weak structure of hadrons, but also "probe" the distributions of isospin, hypercharge, etc. The distributions of the various "charges" need not, generally speaking, coincide.

Unlike leptons, whose structure can be neglected when the deep inelastic processes are studied, in the case of pure hadronic processes the structure of the "probing" hadrons must show up in the presence of (transitional) "form factors" like $\langle c \mid J_\alpha \mid a \rangle$.

Therefore, the functions $F_j(t, M^2)$ in the asymptotic formula for $f_{ab\to cX}$ are factors comprising both the elastic (transitional) form factors of the "probing" particles a and c, and the inelastic form fac-

tors (structure functions) $W_{bj}(\nu, q^2)$ of the target. When ν and $|q^2| \gg m^2$, these structure functions are expected to behave like the structure functions of deep inelastic processes, e.g. manifest an approximate scale invariance.

2.7 SCALE REGULARITIES IN INCLUSIVE PROCESSES

We noted above that the most striking properties of the high energy inclusive cross sections are the various scale regularities.

Certain theoretical arguments for the validity of the scaling for hadron processes have been mentioned in [4], yet the only general conclusion has been the hypothesis of universal scaling nature at high energies.

Many phenomenological models were invented after this to describe and explain, with different degrees of success, the scaling of inclusive spectra for the purely hadronic processes. Below we shall briefly describe the schemes which became widespread. Note at once that up to now the physical reasons for this phenomenon are improperly understood.

The scaling in the deep inelastic processes, at least at the phenomenological level, is much better understood than that in the hadron-hadron processes.

Back in 1964 Markov indicated that leptons could in a way interact with hadrons like point particles. This idea was later defined more specifically in the parton model. If a hadron is assumed to consist, using some sort of kinematics, from point particles (partons). Bjorken scaling follows at once. However, no general argument that can infer from theoretical fundamentals can be made for such behavior. It is only possible to state that Bjorken scaling does not contradict the fundamentals [26].

In 1972 scaling regularities were identified for inclusive processes in the high transverse momenta range but these far exceed what would be expected from extrapolations from the low $k_\perp$ range, where the spectra fall off very fast with $k_\perp$, viz. ($\sim \exp(-Bk_\perp)$).

Yet at high $k_\perp = x_\perp \sqrt{s}/2$ and for fixed angles (in the C.M.S.), experimental data are well described by the formula

$$f_{ab\to cX}(s, k_\perp, \theta) \simeq \frac{1}{k_\perp^N} \varphi(x_\perp, \theta),$$

i.e. the scaling is satisfied for $k_\perp^N f_{ab\to cX}$.

Hereafter we shall detail the concepts which make these scale regularities understandable.

2.7.1 Hadron-Hadron Processes in the Low Transverse Momenta Range

2.7.1a The Automodel Principle

The automodel principle has been put forward in [29]. Every physical parameter that describes particle interactions, say F, is specified in terms of two space scales, a longitudinal one L_z, and a transverse one, $L_\perp$ (in the center-of-inertia system of colliding particles). The dimension of a physical quantity is thus

$$[F] = L_z^m L_\perp^l.$$

The automodel principle is that the dependence of inclusive cross sections on the initial energy is controlled either by a longitudinal or by transverse dimension according to which momentum component of the detected particle grows in proportion to the initial energy.

Consider an inclusive process

$$a + b \to c + X.$$

Provided the transverse momentum is limited, and $k_{zc} \to \infty$, then

$$E_c \simeq |k_{zc}| + \frac{m_c^2 + k_{\perp c}^2}{2\,|k_{zc}|} + \ldots \simeq |k_{zc}|.$$

So here the energy is a longitudinal quantity.

Consider a scale transformation:

$$k_z \to \lambda k_z, \quad \mathbf{k}_\perp \to \mathbf{k}_\perp.$$

According to the automodel principle, the λ-dependence of the physical quantity is controlled by the transverse dimension:

$$F \to \lambda^{-m} F.$$

A single particle inclusive spectrum

$$f_{ab \to cX} = \sum_n n_c 2E_c \frac{d\sigma_{ab}^n}{dk_{zc}\, d\mathbf{k}_{\perp c}} (2\pi)^3$$

has dimensions

$$[f_{ab \to cX}] = L_\perp^4,$$

since it is obvious that

$$[d\sigma_{ab}^n] = L_\perp^2, \quad \left[\frac{dk_{zc}\, d\mathbf{k}_{\perp c}}{E_c}\right] = L_\perp^{-2}.$$

Hence in this case $l = 0$, and the required invariance with respect to transformation (2.38) leads to

$$f_{ab \to cX}(\sqrt{s},\ k_{zc},\ \mathbf{k}_{\perp c}) \to \tilde{f}_{ab \to cX}\left(\frac{2k_{zc}}{\sqrt{s}},\ \mathbf{k}_{\perp c}\right),$$

since the initial energy $\sqrt{s}/2$ is a longitudinal type quantity. A regime in which $s \to \infty$, and $\mathbf{k}_{\perp c}$ and $2k_{zc}/\sqrt{s}$ are kept fixed pertains to the so-called fragmentation region of the initial particles.

If the detected particle momentum is bounded, this corresponds to the pionization or central region. The automodel principle can not be applied here in a straightforward fashion.

2.7.1b Regge Poles

In Sec. 2.5 which was dedicated to an analysis of the $3 \to 3$ analytical properties of the amplitude we mentioned a paper [5] by Mueller, who assumed there to be a relation between the inclusive spectrum and a certain amplitude discontinuity in $3 \to 3$ scattering. True, this fact is not really exploited in [5] where only the inclusive spectrum *per se* is subjected to Regge analysis. Yet in reality, an inclusive spectrum can be treated as a discontinuity in a function which is not the proper $3 \to 3$ transition amplitude, as can be inferred from [23].

Indeed, the results of Sec. 2.5 lead to

$$\widetilde{R}_+ - \widetilde{R}_- = i\,(A_{ac\to bX} - A_{ab\to cX}), \tag{2.39}$$

where

$$\widetilde{R}_+ = \int dy_c\, dy_{c'}\, dy_{b'} e^{ip_c(y_{c'}-y_c)+ip_b y_{b'}} \langle a|\delta_c^+\delta_{b'}^-\delta_{c'}^+ J_b|a\rangle,$$

$$\widetilde{R}_- = \int dy_c\, dy_{c'}\, dy_{b'} e^{ip_c(y_{c'}-y_c)+ip_b y_{b'}} \langle a|\delta_{b'}^-\delta_c^+\delta_{c'}^+ J_b|a\rangle,$$

$$A_{ac\to bX} = \sum_n (2\pi)^4\, \delta(p_a + p_c - p_b - P_{n'})$$

$$\times \left| \int dy_c e^{ip_c y_c} \left\langle a \left| \frac{\delta J_c(y_c)}{\delta\varphi_b(0)} \right| n' \right\rangle \right|^2,$$

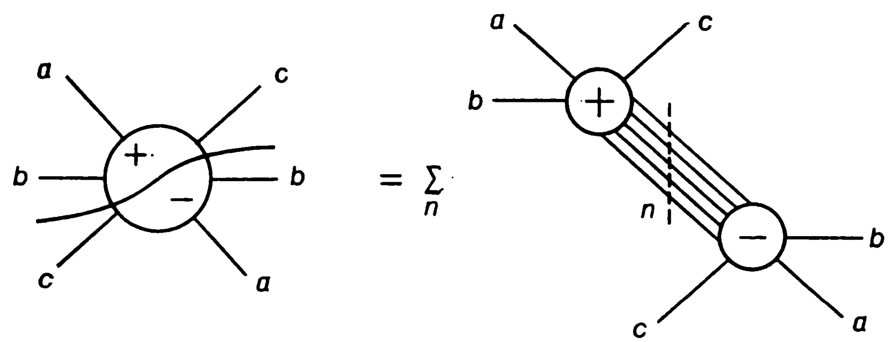
$$A_{ab\to cX} = \sum_n (2\pi)^4\, \delta(p_a + p_b - p_c - P_{n'})$$

$$\times \left| \int dy_b e^{ip_b y_b} \left\langle a \left| \frac{\delta J_b(y_b)}{\delta\varphi_c(0)} \right| n' \right\rangle \right|^2.$$

Then in the domain $q^0 \equiv p_c^0 - p_b^0 > \sqrt{M_+^2 + \mathbf{q}^2} - m_a$ $(\mathbf{p}_a = 0)$, we know $A_{ac\to bX} = I_{ac} f_{ac\to bX}$ remains and in the domain $q_0 < m_a - \sqrt{M_-^2 + q^2}$ we know $A_{ab\to cX} = I_{ab} f_{ab\to cX}$ remains, where $M_\pm$ are the respective minimum masses of the intermediate states, and I_{ac} (I_{ab}) is the invariant flux corresponding to the initial state ac (ab). In some cases there is a gap in the q^0-values, where $\widetilde{R}_+ - \widetilde{R}_- = 0$.

Expression (2.39) is similar to the unitarity condition for the $2 \to 2$ amplitudes yielding the total cross sections of the crossed channels. However, the analytical properties of $\tilde{R}_{\pm}$ have not as yet been sufficiently investigated for an attempt to be made to establish, say, Froissart type bounds.

Thus, it must be kept in mind that even though at the level of diagrams like

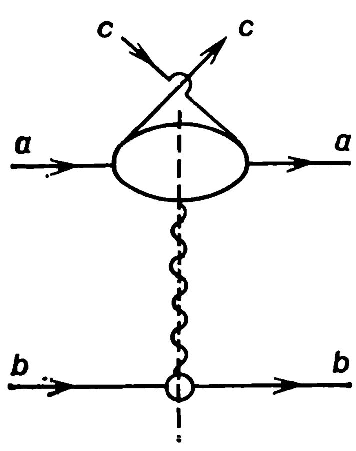


the relation of the $3 \to 3$ amplitude to the inclusive spectra seems, at first sight, exceedingly self-explanatory, in reality we basically know a little about how to understand these relations and what then yields in the sense of inclusive spectra generalities.

Nevertheless, in a purely phenomenological way a naive representation of an inclusive spectrum as being the $3 \to 3$ amplitude discontinuity such as this has proved extremely useful. It has enabled, in part, the Regge approach to be extended up to inclusive spectra.

Roughly speaking, it all amounts to the following. In some kinematic range all the dependence on the "large" parameter s is factorized as a power index of $(s/s_0)^\alpha$, where α is a Regge trajectory that is specified by the quantum numbers in an appropriate "t-channel".

Consider the fragmentation region for the particle a, i.e. $s \to \infty$, $t = (p_a - p_c)^2$ and $x = 2p_{zc}/\sqrt{s} \simeq 1 - M^2/s$ kept fixed, where $M^2 = (p_a + p_b - p_c)^2$. Then the next diagram follows:



where the wavy line denotes a reggeon exchange, i.e. a factor $(M^2)^{\alpha_{b\bar{b}}(0)} \sim s^{\alpha_{b\bar{b}}(0)}$. Since in the t-channel the vacuum quantum numbers "flow", then, assuming the existence of a "vacuum pole (the pomeron)" with $\alpha(0) = 1$, we have

$$f_{ab\to cX} \simeq \frac{1}{s} s^{\alpha(0)} \varphi\left(t, \frac{M^2}{s}\right) = \tilde{\varphi}(\mathbf{p}_{c\perp}, x),$$

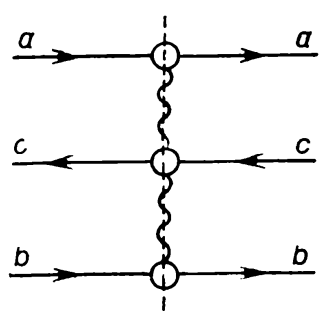
since $t \simeq \frac{p^2_{c\perp}}{x}$ and $M^2 \simeq s(1-x)$.

Thus we obtain the scaling associated with the longitudinal momentum.

In the central domain, where the invariants s, $p_a p_c$ and $p_b p_c$ are large

$$p_a p_c \sim p_b p_c \sim \sqrt{s} \cdot p_{c\perp},$$

a two-reggeon diagram is descriptive:



and the inclusive spectrum has the form

$$f_{ab\to cX} \simeq \frac{1}{s} (p_a p_c)^{\alpha(0)} (p_b p_c)^{\alpha(0)} g(\mathbf{p}_{\perp c}) \sim \tilde{g}(\mathbf{p}_{\perp c}).$$

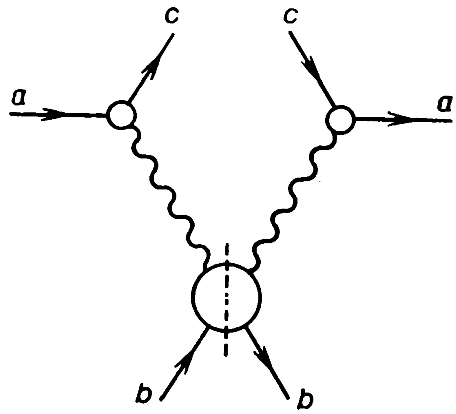
Thus, in the central domain the inclusive spectrum has a finite, energy-independent magnitude which corresponds to the existence of the limit

$$\tilde{\varphi}(\mathbf{p}_{\perp c}, x) \xrightarrow[x\to 0]{} \tilde{g}(\mathbf{p}_\perp).$$

The correction terms responsible for an exchange by the secondary trajectories ($\alpha = 1/2$) have an asymptotics $\sim s^{-1/4}$ (the "RP" term).

In the framework of the Regge scheme the correction sign cannot be established, yet experimental data suggest it is negative. The extinction of such terms as $s \to \infty$ leads to the growth of inclusive spectra in the central domain [29].

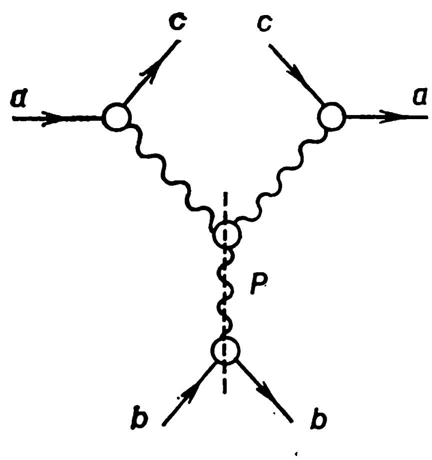
Finally, if we consider the domain as $s \to \infty$; and t and M^2 are kept fixed, the inclusive spectrum is given by the diagram



and is of the form

$$f_{ab\to cX} \sim \frac{1}{s} s^{2\alpha_{a\bar{c}}(t)} \cdot R_{a\bar{c};\, b}(t, M^2).$$

If M^2 is sufficiently large ($m^2 \ll M^2 \ll s$), then a pomeron can be "passed" through the lower block:



and we eventually have

$$f_{ab\to cX} \sim \left(\frac{s}{M^2}\right)^{2\alpha_{a\bar{c}}(t)-1} g(t).$$

This is the so-called "triple reggeon" limit.

So, the scaling takes place in the Regge scheme. Note that in the more sophisticated schemes that allow for branch points in the plane of complex momenta (e.g., the "reggeon" field theory), additional factors of the $(\ln s)^n$ type can originate which violate the scaling [30] in an "actual" way.

2.7.1c Parton Model

The basic postulate of the parton model is the assumption that a hadron is, at the limit of infinite momentum, a set of point-like objects, the partons. The partons' momentum distribution is described by a function $f_H^i(x, \mathbf{k}_\perp)$, where $x = k_{\parallel i}/P_H$ for the frame of reference in which the hadron moves along the z-axis with momentum $P \to \to \infty$. The function $f_H^i(x, \mathbf{k}_\perp)$ is not yet sufficiently clearly defined in terms of the Heisenberg fields and it is not, therefore, possible to elucidate its general properties. In addition, f_H^i is not measurable in a pure form, since it enters physical cross sections in combination with other factors. In a number of cases it plays a principal role.

Historically, the parton model was first invented to explain the scaling in deep inelastic processes. Yet recently the first steps have been made to apply the parton model to hadronic inclusive processes with low $k_\perp$. Certainly, the role of a distribution function here declines appreciably owing to the need to allow for partons to interact. Without this the process could not generally proceed and so a new (and in general unknown) quantity which describes the parton's interaction has been added here. Theory (e.g., QCD) has so far been unable to say anything about quantities which do not involve large virtualities (the necessary condition for applying perturbation theory). Therefore, the application of the parton model to inclusive hadron-hadron processes with low $k_\perp$ was until now kept at a phenomenological level.

Let us briefly describe the appropriate formulas. Suppose a π^+-meson, for example, is created in the fragmentation region during the process $p + p \to \pi^+ + X$. Then the π^+-meson with a large fraction of longitudinal momentum can be treated as having obtained it mainly from the u-quark of one of the initial protons, with a $\bar{d}$-quark carrying off a small fraction of the momentum (this agrees with the data on deep inelastic processes). Whence it follows that

$$\frac{d\langle n_{\pi^+}\rangle}{dx} \sim f_P^u(x). \tag{2.40}$$

The function $f_P^u(x)$ can be taken from the deep inelastic scattering data. The choice of the corresponding values of q^2 is essential and must be low. Such relationships agree quite well with experiments for the processes where the appropriate distribution functions can be found from independent experiments on deep inelastic scattering. This in turn suggests that inclusive processes with low $k_\perp$ can be used to determine the structure functions for hadrons, for which deep inelastic scattering experiments cannot be carried out using modern technology.

For example, when measuring the inclusive spectrum of the process $K^+p \to \pi^+X$ at fairly high x, we might think we could obtain data about the u-quark distribution function in the K^+-meson [31]. True

enough, this possibility appears extremely interesting and, in a sense, even unique. It might, however, be appropriate to point out that in this case (in contrast to processes with high virtualities or transverse momenta) there are no sufficiently reliable and regular computational techniques like perturbation theory, for example.

2.7.2 High Transverse Momenta

2.7.2a Automodelity

Consider the inclusive cross section,

$$\frac{d\sigma^{\text{incl}}_{ab\to cX}}{dk_\perp^2} = \int \frac{dk_\parallel}{16\pi^2 E_c} f_{ab\to cX}(\sqrt{s},\ k_\parallel,\ k_\perp^2).$$

It is evident that

$$\left[\frac{d\sigma^{\text{incl}}_{ab\to cX}}{dk_\perp^2}\right] = L_\perp^4.$$

The automodel principle suggests that at high $k_\perp = x_\perp \sqrt{s}/2$ (i.e. $x_\perp \neq 0$ and kept fixed) "power-like" behavior takes place, viz.

$$\frac{d\sigma^{\text{incl}}_{ab\to cX}}{dk_\perp^2} \simeq \frac{1}{k_\perp^4}\,\varphi(x_\perp),$$

where $\varphi(x_\perp)$ is the unknown function which could contain a weak (logarithmic) dependence on $k_\perp^2$.

2.7.2b Regge Model

The Regge-Mueller expressions for inclusive spectra may formally be considered at high $k_\perp$, too, for example, in the central domain. However, within the framework of Regge approach there is no information about the dependence of the vertices that connect both the reggeons and particles on their arguments. In other domains (fragmentation) we are also uncertain about the Regge trajectories, $\alpha(t)$ (as well as the residues) at $t \to -\infty$. The leading trajectories at low t may turn out to be corrective at high t range. All attempts to include high t's into reggistics are based on some assumptions concerning the behavior of the trajectories and the residues in this case. It appears that the "language" of Regge trajectories ceases to be adequate, physically at high t, since the character of interactions here (i.e. at small distances) differs in principle from the picture for low t (i.e. at large distances).

2.7.2c Parton Model

An inclusive process is described in the parton model thus: two partons, one from each of the incident hadrons, collide and scatter through a large angle [32]. A subsequent fragmentation of the scattered partons yields a hadron with a high transverse momentum. Eventually, the scattered partons transform into two hadronic jets that form large angles with respect to the primary beam. Thus, the inclusive spectrum is composed of three components: the distribution functions of partons in the initial hadrons, the parton's scattering amplitude, and the functions governing the fragmentation of the partons into the detected hadrons.

Some independent information about the distribution functions and fragmentation functions can be inferred from the data on deep inelastic scattering and inclusive e^+e^- annihilations. If we want to predict the evolution of the inclusive spectra, the parton-parton scattering amplitude must be taken from theory. Commonly the role of the partons is played by quarks and gluons. However, it can also be played by virtual hadrons.

Consider a scattering amplitude in the general case in which each of the initial and final particles can contain a certain number of elementary components, "constituents". In particular, these particles can themselves be the constituents.

Now assume that the amplitude of a transition $\langle m \mid T \mid n \rangle$, where m and n are the numbers of the constituents in the initial and the final states, obeys the automodel principle in the domain in which all the invariant constituents from the different particles are large, i.e. the large angle domain. Then, according to the automodel principle, the dependence of $\langle m \mid T \mid n \rangle$ on the energy $\sqrt{s}$ is controlled by its physical dimension. Since

$$\langle \mathbf{k} \mid \mathbf{k}' \rangle = (2\pi)^3 2E\delta(\mathbf{k} - \mathbf{k}'),$$

$[\langle m \mid] = \mu^{-m}$, μ being an arbitrary mass unit. On the other hand, the relation

$$S_{\omega\alpha} = \langle \omega \mid \alpha \rangle + \ldots + i(2\pi)^4 \delta(P_\omega - P_\alpha) \langle \omega \mid T \mid \alpha \rangle$$

suggests that

$$[T] = \mu^4.$$

Whence

$$[\langle m | T | n \rangle] = \mu^{4-m-n}.$$

This means that

$$\langle m | T | n \rangle \simeq (\sqrt{s})^{4-m-n} \Phi(\theta_i),$$

where the θ_i are dimensionless parameters (angles). Whence we obtain the famous "quark counting" formula [33] for the scattering

$$A + B \rightarrow C + D$$

at large angles, viz.

$$\left(\frac{d\sigma}{dt}\right)_{AB\rightarrow CD} \approx \frac{1}{s^{n_A+n_B+n_C+n_D-2}} \varphi(\theta), \tag{2.41}$$

where $n_{A,B,C,D}$ are the numbers of constituents in the particles A, B, C, D.

The asymptotic law (2.41) has also been obtained independently in [34] by analyzing Feynman diagrams. Experimental data agree quite well with the quark counting rule.

In (2.41) we took into account only the "valence" quarks. If configurations that contain "sea" quarks in addition to valence quarks are also taken into account, then clearly corrections which fall off more rapidly with s will be produced.

Let us return to the inclusive spectra. According to the parton model, the inclusive spectrum in the high $k_\perp$-range is of the form

$$f_{AB\rightarrow CX} = \sum_{a,\,b,\,c,\,d} \int dx_a\, dx_b f_A^a(x_a)\, f_B^b(x_b) \times \left(\frac{d\sigma}{d\hat{t}}\right)_{ab\rightarrow cd} \frac{1}{z_C} \bar{D}_c^C(z_C), \tag{2.42}$$

where $f_A^a(x_a)$ is the distribution function for parton a in particle A over longitudinal momentum fraction, $x_a = k_{\|a}/P_A$; $\bar{D}_c^C$ is the $c \rightarrow C$ fragmentation function; $z_C = P_{\|C}/k_{\|c}$, and $\left(\frac{d\sigma}{d\hat{t}}\right)_{ab\rightarrow cd}$ describes the elementary subprocess $a + b \rightarrow c + d$, with

$$\hat{t} = (p_a - p_c)^2 \simeq \frac{x_a}{z_C}(p_A - p_C)^2 \simeq \frac{2x_a}{z_C} \frac{(\sqrt{x^2+x_\perp^2} - x)}{x_\perp^2} p_{\perp C}^2;$$

$$x_\perp = \frac{2P_{\perp C}}{\sqrt{s}}, \quad x = \frac{2P_{\|C}}{\sqrt{s}}.$$

From formula (2.42) it is not difficult to see that the function $P_{\perp c}$ is governed by the cross section of the elementary subprocess $a + b \rightarrow c + d$.

If $a, \ldots, d$ are quarks, then from formula (2.42) we can obtain that

$$f_{AB\rightarrow CX} \simeq \frac{1}{P_{\perp C}^4} \Phi(x, x_\perp).$$

If, for example, $a = c = q$ (the quark) and $b = d = M$ (the meson), then the appropriate contribution to the inclusive spectrum is given, according to (2.41), by the expression

$$\frac{1}{p_{\perp C}^{8}} \tilde{\Phi}(x, x_{\perp}).$$

Taking into account all possible contributions, we obtain the following general formula:

$$f_{AB \to CX} \simeq \sum_i \frac{1}{p_{\perp C}^{N_i}} \Phi_i(x, x_{\perp}).$$

The known distribution and fragmentation functions enable us to discover definite information about the behavior of $\Phi_i(x, x_{\perp})$. It can be proved that the different terms of the sum dominate in different ranges of x and $x_{\perp}$. A detailed study of these more intricate questions was carried out in [35].

2.7.3 Deep Inelastic Scattering

2.7.3a The Automodel Principle

According to the automodel principle, the structure functions of deep inelastic scattering as $-q^2$ and $\nu \to \infty$ (the ratio $x = -q^2/\nu$ kept fixed) change in the scale transformations,

$$q \to \lambda q \text{ and } p \to \lambda p,$$

as homogeneous functions of the appropriate dimension. Since

$$[W_1] = \mu^0, \quad [W_2] = \mu^{-2},$$

then

$$W_1(\lambda q, \lambda p) = W_1(q, p)$$

and

$$W_2(\lambda q, \lambda p) = \lambda^{-2} W_2(q, p).$$

Whence in the Bjorken domain

$$W_1(q^2, \nu) \to F_1(x) \text{ and } \nu W_2(q^2, \nu) \to F_2(x).$$

In this connection Bogolyubov has suggested an interesting hypothesis about a possible analogy between deep inelastic processes and the dynamics of a point explosion in gas dynamics.

From the physical standpoint, this suggests that in the Bjorken domain a hadron's size, when is summed up over all channels, ceases to be essential and dimensional factors (masses) disappear from the analysis and we arrive at the scaling.

If, however, we try to substantiate this inside the framework of quantum field theory, we discover that, apart from the physically

uninteresting possibility of $g\varphi_4^3$, there is an effective mass parameter in any renormalized field theory due to the higher approximations and renormalizations, and the resultant logarithms yield anomalous dimensions which violate the scaling.

2.7.3b Regge Poles

Certain additional assumptions allow us to make some definite conclusions about structure functions of deep inelastic scattering using the Regge approach.

In fact, a standard analysis of the imaginary part of the amplitude of a virtual photon scattered by a hadron yields the following expression for the cosine of the scattering (non-physical) angle in the t-channel:

$$\cos\theta_t \sim \frac{\nu}{m\sqrt{q^2}}.$$

The general expression for the structure function W_1 is

$$W_1(\nu, q^2) \simeq \sum_i b_i(q^2)\left(\frac{\nu}{m\sqrt{q^2}}\right)^{\alpha_i(0)},$$

where the α_i are Regge trajectories, viz. P (the pomeron), ρ, ω, ...

Scale invariance in the Bjorken domain can be achieved once it is assumed that

$$b_i(q^2)_{-q^2\to\infty} \sim \left(\frac{1}{\sqrt{q^2}}\right)^{\alpha_i(0)} \bar{b}_i.$$

Then

$$W_1(x, q^2) \simeq \bar{b}_P\frac{1}{x} + \bar{b}_R\frac{1}{\sqrt{x}} + \ldots,$$

where the first term pertains to an exchange of a pomeron ($\alpha = 1$), and the second one to the meson trajectories ($\alpha \simeq 1/2$). This behavior qualitatively reflects the basic experimental data for low x.

The conjecture about the q^2-asymptote residues cannot be justified within the Regge approach. Besides, at high x the Regge formulas become absolutely invalid.

2.7.3c Parton Model

The cross section of the deep inelastic process

$$e + p \to e' + X$$

has the form

$$\frac{d\sigma}{dE'\,d\Omega'} = \frac{4\alpha^2E'^2}{mQ^4}\left[2\sin^2\frac{\theta}{2}W_1(q^2, \nu) + m^2\cos^2\frac{\theta}{2}W_2(q^2, \nu)\right],$$

where E' is the final-electron energy, θ the scattering angle of the final electron in the laboratory frame ($\mathbf{p} = 0$), $\nu = 2pq$, and $q = k_e - k_e'$. The structure functions W_1 and W_2 enter the tensor $W_{\mu\nu}$ thus:

$$W_{\mu\nu} = \left(-g_{\mu\nu} + \frac{q_\mu q_\nu}{q^2}\right) W_1$$
$$+ \left(p_\mu - \frac{qp}{q^2} q_\mu\right)\left(p_\nu - \frac{qp}{q^2} q_\nu\right) W_2.$$

Compute W_1 and W_2 in the parton model, where the partons have spin 1/2 and are sorted by the index i:

$$W_{\mu\nu} = \sum_i e_i^2 \int \frac{dz}{2z} d\mathbf{k}_\perp f_i(z, \mathbf{k}_\perp)$$

$$\times \operatorname{Tr}[\hat{k}\gamma_\mu(\hat{k}+\hat{q})\gamma_\nu] 2\pi\delta(q^2 - 1qk).$$

Assuming the integral converges in $\mathbf{k}_\perp$, we obtain

$$2kq \simeq z\nu,$$
$$W_1(x, q^2) = \sum_i e_i^2 \int d\mathbf{k}_\perp f_i(x, \mathbf{k}_\perp),$$
$$W_2(x, q^2) = \sum_i e_i^2 \frac{4x}{\nu} \int d\mathbf{k}_\perp f_i(x, \mathbf{k}_\perp).$$

Thus, we obtain the Bjorken scale invariance:

$$W_1(x, q^2) \to F_1(x),$$
$$\nu W_2(x, q^2) \to F_2(x),$$

with

$$F_2(x) = 4xF_1(x) \tag{2.43}$$

Whence it follows, in particular, that for deep inelastic processes we obtain straightforward information about nucleonic structure.

Had we considered scalar partons, we would have $W_1 = 0$. Equation (2.43), which is known as the Callan-Gross relation, is satisfied in experiments fairly accurately. This can be regarded as evidence in favor of the fermionic charged partons.

There is no way of computing the distribution functions using the parton model, i.e. the x-dependence must come from experimental data and then used for the other processes such as the production of high $\mathbf{k}_\perp$ hadrons.

Nevertheless, by making another important assumption, viz. that the charged partons are quarks, a number of useful general relationships, that can be checked in experiment, can be obtained.

Actually, we have for the proton and neutron structure functions

$$W_{1p}(x) = \sum_{i=u,\, d,\, s} e_i^2 f_p^i(x) = \frac{4}{9}(u(x) + \bar{u}(x))$$

$$+\frac{1}{9}(d(x) + \bar{d}(x)) + \frac{1}{9}(s(x) + \bar{s}(x)),$$

$$W_{1n}(x) = \sum_{i=u,\, d,\, s} e_i^2 f_n^i(x) = \frac{4}{9}(d(x) + \bar{d}(x))$$

$$+\frac{1}{9}(u(x) + \bar{u}(x)) + \frac{1}{9}(s(x) + \bar{s}(x)).$$

A proton consists of three valence quarks (uud) and of an arbitrary number of the "sea" quarks, i.e. the pairs $u\bar{u}$, $d\bar{d}$, $s\bar{s}, \ldots$. Likewise, a neutron is composed of the valence (udd) and sea ($u\bar{u}$, $d\bar{d}$, $s\bar{s}$, ...) quarks. Provided the "sea" of $SU(3)$ is symmetric, we have

$$\int_0^1 dx\, [W_{1p}(x) - W_{1n}(x)] \simeq \frac{1}{3} \int_0^1 dx\, [u(x) - d(x)].$$

Since the protons contain one more u-quark than it does d-quarks, then

$$\int_0^1 dx\, [u(x) - d(x)] = \bar{N}_u - \bar{N}_d = 1,$$

and, as a result,

$$\int_0^1 dx\, [W_{1p}(x) - W_{1n}(x)] \cong \frac{1}{3}.$$

2.7.4 Scaling Violation

The scaling laws discussed in the foregoing sections are approximate. In fact, functions of dimensionless momenta ratios manifest a "residual" energy dependence. This fact can be interpreted in different ways:

(1) the inclusive cross sections really do tend to the scale invariant limits, and the energy dependence we observe is due only to correction terms whose contribution is infinitesimal at fairly high energies;

(2) the scale invariance really is violated, i.e. the energy dependence cannot be neglected at any energy.

Besides, the quantities measured can be subdivided into two classes according to case (1) and (2), respectively.

Consider several examples.

1. *Hadron-Hadron Processes*. In a "naive" Regge scheme, as seen above, a true scaling occurs for the inclusive spectra, viz.

$$f_{ab\to cX}(s, x, \mathbf{k}_\perp) \to \tilde{\varphi}_c(x, \mathbf{k}_\perp),$$

of the process $a + b \to c + X$, with $\tilde{\varphi}_c(x, \mathbf{k}_\perp)$ having a finite limit in the central domain ($x \to 0$). The consequence is that the mean multiplicity growing logarithmically, viz.

$$\langle n_c \rangle \sim \ln s.$$

Thus, naive reggistics implements situation (1). A deviation from the scaling observed is interpreted in terms of a contribution of secondary trajectories, which disappears as $s \to \infty$.

Alternatively, in the Cheng-Wu model [36], spectra grow power-like in the central domain and, consequently, we have

$$\langle n_c \rangle \sim \frac{s^\alpha}{(\ln s)^\beta}; \quad \alpha = \frac{1}{2} - \varepsilon, \quad \beta = 2\varepsilon,$$
$$0 < \varepsilon < 1/2.$$

Obviously, we have here situation (2).

2. *Deep inelastic processes*. In the standard parton model [37] the structure functions of deep inelastic lepton-hadron scattering tend to functions of $x = -q^2/\nu$. The corrections related to the masses of hadrons and leptons and the transverse motion of the partons fall off at $\sim 1/q^2$.

A parton picture using a perturbative QGD holds with violated scaling [38]. Despite the effective interaction is weakened at short range, i.e.

$$\alpha_s(q^2) \sim [\ln(-q^2/\Lambda^2)]^{-1},$$

this weakening is insufficient to compensate for the high probability that soft and collinear gluons can be emitted and the probability leads to factors $\sim \ln(-q^2/\mu^2)$. This is the reason why deep inelastic structure functions cannot be computed by limiting to a finite number of terms of the effective coupling constant $\alpha_s(q^2)$ in the perturbative series, i.e. the whole series must be summed. This works successfully using the leading logarithmic approximation detailed in [39] for assymptotically nonfree theories and can then be extended to cover QCD. As a result, the structure functions become q^2-dependent, although for a large fraction of x values (which are not close to $x = 0$ and $x = 1$) this dependence is relatively weak, i.e. the specific parameter is $\sim \ln(\ln(-q^2))$. An analogous situation is true for other inclusive processes, i.e.

$$e^+e^- \to hX, \; pp \to \mu^+\mu^- X, \; h_1h_2 \to h_3(k_\perp)X, \; \ldots .$$

In some processes, however, the entire energy dependence of the cross sections is governed by the coupling constant $\alpha_s(q^2)$. These are the so-called infrared stable processes, e.g. $e^+e^- \to$ hadrons, angular distributions of hadron jets in the e^+e^--annihilation in the cones with fixed openings, etc. As $q^2 \to \infty$ the cross sections of such processes tend to the relations defined by the naive parton model, and the corrections are then found from the theory of perturbations in $\alpha_s(q^2)$ [40] and, certainly, they fall off.

Yet another source of scaling violation is to allow for non-perturbative fluctuations of the QGD vacuum, that is the instantons [41] and the fluctuations related to the non-zero average quark and gluon fields, $\langle 0 \mid \bar{q}q \mid 0 \rangle$ and $\langle 0 \mid \mathbf{F}_{\mu\nu}\mathbf{F}^{\mu\nu} \mid 0 \rangle$, respectively [42]. However, their contributions disappear as inverse power of q^2.

Thus, the inclusive cross sections of the processes in the QCD which are not infrared-stable must actually violate the scaling at high q^2.

Unfortunately, the cause for the scaling violation in "soft" hadron-hadron inclusive processes is inaccessible for a QCD analysis, since it requires, perhaps, a more substantial extension of the theory beyond the framework of perturbation theory and therefore is more closely associated with the confinement problem.

Appendix

On the Analytical Properties of the Amplitude of the Process

$$a + b \to n_c c + n_d d + \ldots \qquad (*)$$

in the Angular Variables that Define the Direction of the Unit Vector

$\mathbf{n} = \mathbf{k}_a / |\mathbf{k}_a|$ in the C.M.S. $\mathbf{k}_a + \mathbf{k}_b = 0$ [11]

In Sec. 2.3 it was shown that the amplitude of process (*), $T^n_{ab}(s, \mathbf{n})$, could be presented in the form

$$T^n_{ab}(s, \mathbf{n}, \ldots) = \int\limits_{x_L(s)}^{\infty} dx \int d\mathbf{e} \, \frac{\Phi(x, \mathbf{e}; s, \ldots)}{x - \mathbf{e}\cdot\mathbf{n}}, \qquad (**)$$

where $x_L(s)$ is the semi-major axis of the Lehmann ellipse.

Consider the analytic properties of the integral on the right-hand side of (**). For the complex vector $\mathbf{n} = \mathbf{x} + i\mathbf{y}$ the integral is seen to be continued into domain

$$B(s) = \{\mathbf{n} \mid |\mathbf{x}| < x_L(s), \ \mathbf{y} \in R^3\}.$$

Since the amplitude T^n_{ab} is not initially given at all $\mathbf{n}$ values, but only for the unit sphere, $\mathbf{n}^2 = 1$, then this continuation is not, gener-

ally speaking, unique, i.e. various functions of the complex vector $\mathbf{n}$ could coincide in the subset $|\mathbf{x}| = 1$, $\mathbf{y} = 0$. A unique continuation would be the one onto the complex sphere

$$\mathbf{n}^2 = (\mathbf{x} + i\mathbf{y})^2 = 1.$$

Thus, the amplitude $\overset{*}{T}{}^n_{ab}(s, \mathbf{n}, \ldots)$ analytically continues into the domain $\mathscr{D}(s)$ on a complex sphere

$$\mathscr{D}(s) = \{\mathbf{n} \mid \mathbf{n}^2 = 1, \quad \mathbf{n} \in B(s)\},$$

and this continuation is unique.

Likewise, the function $T^n_{ab}(s, \mathbf{n}, \ldots)$ can be shown to continue analytically into the domain $\mathscr{D}(s)$.

Whence it follows that the differential cross section

$$\frac{d\sigma^n_{ab}}{d\mathbf{n}} = \frac{1}{s} \int d\xi_n \overset{*}{T}{}^n_{ab}(s, \mathbf{n}, \xi_n)\, T^n_{ab}(s, \mathbf{n}, \xi_n)$$

also analytically continues into the domain $\mathscr{D}(s)$ which is independent of the parameters ξ_n.

In order to pass over to angular coordinates, we can choose the following parametrization of the vector $\mathbf{n}$:

$$\begin{aligned} n_1 &= \sqrt{1 - z^2}\, \frac{1}{2}(w + 1/w), \\ n_2 &= \sqrt{1 - z^2}\, \frac{1}{2i}(w - 1/w), \\ n_3 &= z. \end{aligned} \tag{***}$$

When $w = e^{i\varphi}$, $z = \cos\theta$ and $0 \leqslant \varphi \leqslant 2\pi$, $0 \leqslant \theta \leqslant \pi$ we arrive at the standard parametrization in terms of spherical angles.

The above parametrization of $\mathbf{n}$ is ambiguous in the vicinity of $n_3 = \pm 1$. To avoid ambiguity, the cuts $(-\infty, -1]$, $[1, +\infty)$ must be made.

Thus parametrization (***) is not defined on the entire sphere, but rather in the domain

$$\mathcal{U} = \{\mathbf{n} \mid \mathbf{n}^2 = 1, \quad n_3 \bar{\in} (-\infty, -1] \cup [1, +\infty)\}.$$

This reflects the property of complex analytical manifolds, which include the sphere $\mathbf{n}^2 = 1$, that unified coordinates which are everywhere regular generally fail to be defined on these manifolds. A regular branch $\sqrt{1 - z^2}$ can be isolated in the z-plane with the cuts $(-\infty, -1]$ and $[1, +\infty)$ by means of the conditions $\sqrt{1 - z^2} > 0$ for $-1 < z < 1$. From the conditions pertaining to the domain $\mathscr{D}(s)$, namely,

$$\mathbf{x}^2 - \mathbf{y}^2 = 1, \; \mathbf{x} \cdot \mathbf{y} = 0$$

and

$$\mathbf{x}^2 < x_L^2(s),$$

the analyticity domain in z and w,

$$\left(|1+z||w|+|1-z|\frac{1}{|w|}\right)\times \\ \times\left(|1+z|\frac{1}{|w|}+|1-z||w|\right)<4x_L^2(s),$$
$$z\bar{\in}(-\infty,-1]\cup[1,+\infty),$$

can now be obtained.

Now, let us show that the differential cross section ($d\mathbf{n} = d\cos\theta\ d\varphi$),

$$\frac{d\sigma_{ab}^n}{d\cos\theta}=\int_0^{2\pi} d\varphi\,\frac{d\sigma_{ab}^n}{d\cos\theta\, d\varphi}$$

continues analytically into a Lehmann ellipse without the cuts. To do this, consider the function $(d\sigma_{ab}^n/d\mathbf{n})(s,\mathbf{n})$ in the vicinity of the point $\mathbf{n}=(0,0,1)$. The regular coordinates in this vicinity are

$$z^{\pm}=n_1\pm in_2.$$

The quantity $d\sigma_{ab}^n/d\mathbf{n}$, being an analytic function in $\mathscr{D}(s)$, will be given in the z-coordinates in the vicinity of the point $(z^+, z^-) = (0, 0)$ by the series

$$\frac{d\sigma_{ab}^n}{d\mathbf{n}}=\sum_{ml}a_{ml}(s)\,(z^+)^m\,(z^-)^l.$$

Expressing the new coordinates in terms of the old ones, we obtain

$$\frac{d\sigma_{ab}^n}{d\mathbf{n}}=\sum_{ml}(1-z^2)^{\frac{m+l}{2}}w^{m-l}a_{ml}(s).$$

Assuming now that $w=e^{i\varphi}$ and integrating the series over φ from 0 to 2π, we obtain

$$\int_0^{2\pi}\frac{d\sigma_{ab}^n}{d\mathbf{n}}\,d\varphi=\sum_{m=0}^{\infty}(1-z^2)^m\,a_{mm}(s).$$

Thus, the points ± 1 for the function

$$\frac{d\sigma_{ab}^n}{d\cos\theta}(s,\cos\theta)$$

are not singular and this proves its analyticity in the Lehmann domain without cuts.

References

1. A. A. Logunov, M. A. Mestvirishvili, and Nguyen Van Hieu, Preprint IFVE 67-49K, Serpukhov (1967);
Proc. 1967 Intern. Conf. Part. Fields, New York (1967);
Phys. Lett., **25B**, 611 (1967).
A. A. Logunov, Nguyen Van Hieu, *Lectures at the Intern. School High Energ. Phys.*, Popradske Pleso, Czechoslovakia (1967).
2. A. A. Logunov and Nguyen Van Hieu, *Topical Conf. High Ener. Coll. Hadrons*, v. 2, CERN, Geneva (1968), p. 74.
G. Tiktopoulos and S. B. Treiman, *Phys. Rev.*, **167**, 1408 (1967).
A. A. Logunov, Nguyen Van Hieu, and O. A. Khrustalyov, in: *Problemy Teoreticheskoi Fiziki*, Moscow (1969), p. 90 (in Russian).
3. Yu. B. Bushnin *et al.*, *Phys. Lett.*, **29B**, 48 (1969).
4. T. Benecke *et al.*, *Phys. Rev.*, **188**, No. 5, 2159 (1969);
R. P. Feynman, *Phys. Rev. Lett.*, **23**, 1072 (1969).
5. A. H. Mueller, *Phys. Rev.*, **D2**, 2963 (1970).
6. A. A. Logunov, M. A. Mestvirishvili, and V. A. Petrov, *Ann. Phys.*, **114**, 46 (1978).
7. A. A. Logunov, M. A. Mestvirishvili, and V. A. Petrov, Preprint IFVE 74-66, Serpukhov (1974).
8. E. D. Bloom *et al.*, *Phys. Rev. Lett.*, **23**, 930 (1969).
9. M. Banner *et al.*, *Phys. Lett.*, **41B**, 547 (1972);
F. W. Büsser *et al.*, *Phys. Lett.*, **46B**, 471 (1973).
10. N. N. Bogolyubov, B. V. Medvedev, and M. K. Polivanov, *Problems of the Theory of Dispersion Relations*, Princeton (1955).
11. V. V. Ezhela, Preprint IFVE 71-36, Serpukhov (1971).
12. P. Ascoli and A. Minguzzi, *Phys. Rev.*, **118**, 1435 (1960).
13. V. V. Ezhela, A. A. Logunov, and M. A. Mestvirishvili, *Theor. and Math. Phys.*, 6 (1971).
14. V. V. Ezhela, A. A. Logunov, M. A. Mestvirishvili, and V. A. Petrov, *Theor. and Math. Phys.*, 15, No. 2, 427 (1973).
15. A. Martin, *Scattering Theory: Unitarity, Analyticity, and Crossing*, Springer Verlag, Berlin-New York (1969).
16. G. Tiktopoulos and G. Treiman, *Phys. Rev.*, D6, 2045 (1972).
17. A. A. Logunov, M. A. Mestvirishvili, and V. A. Petrov, *Theor. and Math. Phys.*, **21**, No. 3, 1160 (1974).
18. A. A. Logunov and M. A. Mestvirishvili, *Theor. and Math. Phys.*, **11**, No. 2, 451 (1972).
19. G. L. Rcheulishvili, Preprint IFVE 72-96, Serpukhov (1972).
20. V. Yu. D'yakonov and V. E. Rochev, *Theor. and Math. Phys.*, **32**, No. 1, 644 (1977).
21. V. V. Ezhela and V. A. Petrov, Preprint IFVE 72-51, Serpukhov (1972).
22. G. Giacomelli and M. Jacob, *Phys. Rep.*, **55**, No. 1 (1979).
23. A. A. Logunov *et al.*, *Theor. and Math. Phys.*, **33**, No. 2, 936 (1977).
24. R. Jost and H. Lehmann, *Nuovo Cim.*, **5**, 1598 (1957);
F. Dyson, *Phys. Rev.*, **110**, 1460 (1958).
25. N. N. Bogolyubov, V. S. Vladimirov, and A. N. Tavkhelidze, *Theor. and Math. Phys.*, **12**, No. 1, 619 (1972).
26. B. I. Zav'alov, *Theor. and Math. Phys.*, **17**, No. 2, 1074 (1973);
ibid. 33, No. 3, 1040 (1977);
E. Wieczorek, V. A. Matveev, D. Robaschik, and A. N. Tavkhelidze, *Theor. and Math. Phys.*, **16**, No. 3, 861 (1973).
27. A. A. Logunov, M. A. Mestvirishvili, and V. A. Petrov, *Theor. and Math. Phys.*, **32**, No. 1, 559 (1977).
28. A. K. Likhoded and P. V. Shlyapnikov, *Sov. Phys. Uspekhi*, **21**, No. 1, 1 (1978).
29. V. A. Matveev, R. M. Muradyan, and A. N. Tavkhelidze, *Elementary Particles and Atomic Nucleus*, **2**, No. 1, Moscow (1971).

30. M. S. Dubovikov and K. A. Ter-Martirosyan, *Sov. Phys. JETP*, (1977).
31. P. V. Chliapnikov, V. G. Kartvelishvili, V. V. Kniazev, and A. K. Likhoded, *Nucl. Phys.*, **B148**, 400 (1978).
32. M. Berman, J. D. Bjorken, and J. Kogut, *Phys. Rev.*, **D4**, 3388 (1971).
33. V. A. Matveev, R. M. Muradyan, and A. N. Tavkhelidze, *Nuovo Cim. Lett.*, **7**, 719 (1973).
34. S. Brodsky and G. Farrar, *Phys. Rev. Lett.*, **31**, 1153 (1973).
35. R. Blankenbecker and S. Brodsky, *Phys. Rev.*, D**10**, 2973 (1974).
36. H. Cheng and T. T. Wu, *Phys. Lett.*, **45B**, 367 (1973).
37. R. P. Feynman, *Photon-Hadron Interactions*, Benjamin, Reading (1972). J. D. Bjorken and E. Paschos, *Phys. Rev.*, **185**, 1965 (1969).
38. A. V. Efremov and V. A. Radyushkin, *Rivista di I. Nuovo Cim.*, **3**, No. 2 (1980).
39. V. N. Gribov and L. N. Lipatov, *Yadernaya fizika*, **15**, 781, 1218 (1972).
40. K. G. Chetyrkin, A. L. Kataev, and V. F. Tkachev, *Phys. Lett.*, **85B**, 277 (1979).
41. A. A. Belavin, A. M. Polyakov, A. S. Schwarz, and Y. S. Tyupkin, *Phys. Lett.*, **59B**, 85 (1975).
42. M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, *Nucl. Phys.*, **B147**, 385, 448, 519 (1979).

3 Colored Quarks

N. N. Bogolyubov, V. A. Matveev, and A. N. Tavkhelidze

3.1 INTRODUCTION

A concept of colored quarks as fundamental fermions possessing a specific quantum number, color, and like leptons, being elementary constituents of matter, forms the foundation of current theoretical ideas about the world of elementary particles and atomic nuclei.

When in 1964 Gell-Mann [1] and Zweig [2] suggested the existence of quarks as hypothetical particles that make up mesons and baryons (i.e. every strongly interacting particles observed) quarks were imagined as purely mathematical objects. Using them, the then discovered unitary *SU* (3) symmetry of strong interactions [3] could be described in the most simple and elegant way. These particles, which possess fractional electric charges and are not observable in free state, eventually gained an appropriate physical interpretation.

First of all, constructing hadrons from quarks having spin 1/2 lead to a contradiction with the Pauli principle for systems of particles with half integer spin.

The problem of statistics was not, however, the only obstacle standing in front of the theory. The following questions remain unanswered: why in nature do only those systems which correspond to three quarks or quark-antiquark pairs exist? Why are there no indications of the existence of other multiquark systems?

Of special interest is the question as to whether quarks could exist in the free state (the problem of quark confinement or non-emission).

In 1965 an analysis of these issues led Bogolyubov, Struminsky, and Tavkhelidze [4], as well as Nambu and Han [5], Freund [6], and Miyamoto [7] to the fundamental idea that the quarks possess a new and hitherto unknown quantum number afterwards called *color* [8].

For more than 15 years this idea lies at the base of elementary particle physics. Once hadron spectroscopy problems could be solved, the colored quarks hypothesis led on to the development of the quantum chromodynamics (QCD) of strong interactions, and enlivened numerous versions of the "grand unification" theory (GUT).

The present paper considers the main topics for the development of the theory of colored quarks and exposes a number of the important achievements for the physics of elementary particles, atomic

nuclei, and high energies [9-11] which were obtained from this theory.

The paper begins with a discussion on the dynamic treatment of hadrons as composite quark systems and the construction of form factors and amplitudes of various processes, in which hadrons take part (Sec. 3.2).

A dynamic quark model [4, 12] proposed in 1964 at JINR (Dubna) offered a systematic description of both static and observed parameters of elementary particles (magnetic moments, axial-vector constants of weak transitions, etc.), and hadron form factors [13]. These studies gave an impetus to the development of the current quark models of elementary particles, among which the quark bag model [14, 15] and quark parton model are the most popular.

An important step for the developing the dynamic hadron theory was made by Nambu, who first introduced vector fields, i.e. the carriers of the color interaction, which became the prototype of the QCD gluon field [16]. QCD, whose rapid progress we have observed in the last few years [17], originated then as a result of a unification of the hypothesis of colored quarks and the color SU^c (3) symmetry with the principle of local gauge invariance by Yang and Mills [18].

It is important to emphasize in this respect that the Greenberg hypothesis of the parafermi statistics of quarks [19-20], as shown in Sec. 3.3, does not admit the gauge SU^c (3) symmetry, since this symmetry forms the basis of QCD and the Greenberg hypothesis is thus a physically unacceptable alternative to the hypothesis of colored fermi quarks [21].

It is obviously not possible in a paper of this size to elucidate all the achievements of QCD, the development of which was a substantial advance for the theory of strong interactions.

Secs. 3.4 and 3.5 show how QCD and the ideas of composite hadrons can describe a wide range of phenomena from approximate scale invariance to automodel behavior in a consistent theoretical way and can substantiate the quark counting method for high momenta transfer processes.

The scale invariance discovered when inclusive and deep inelastic processes were investigated is one of the most universal regularities in high energy physics [22].

The automodel hypothesis formulated in [23] results in unified approach to the study of the scale properties of high-energy, strong, weak, and electromagnetic interactions based on similarity principles and dimensional analysis. The compatibility of the automodel hypothesis asymptotic behavior with the fundamentals of quantum field theory has been rigorously corroborated in [24], where a one-to-one relation has been found between the automodel's amplitude and cross-sectional asymptotes for deep inelastic processes and the behavior of a local currents products near the light cone.

In 1973 the automodel hypothesis and the ideas about hadrons' quark structure led to the formulation of the quark counting rules. These specify the pattern of the asymptotic behavior of amplitudes and cross sections of the various exclusive processes depending on a degree of "complexity" of the partaking hadrons [25, 26].

The quark counting formulas describe the numerous experimental data on the elementary particle scattering surprisingly well and enable straightforward information about the number of elementary hadron constituents to be taken from the experiments.

In the last few years the idea of colored quarks and fundamental QCD forces has started penetrating into the theory of nuclear reactions.

It is to be noted that the most immediate manifestation of the quark nuclear structure is an experimentally observed law of the exponential falling of the electromagnetic form factor of deuteron at high momenta transfers, which agrees well with the quark counting formula and exhibits the presence of a hard 6-quark deuteron structure [27, 28].

Sec. 3.6 treats the problems of allowing for the quark degrees of freedom in describing pure nuclear phenomena, especially those taking place at high energies and momenta transfers. A possibility is, in particular, indicated of exciting the "hidden" color in nuclear matter and of a number of other consequences [29].

The final Sec. 3.7 is dedicated to discussing the unified gauge theories of strong and electromagnetic interactions having a spontaneously broken color symmetry and integral charge quarks [30, 31].

3.2 COLORED QUARKS AND HADRON DYNAMICS

3.2.1 A Hypothesis of Colored Quarks

According to the colored quarks hypothesis formulated for the first time in [4-7], quarks obey Fermi-Dirac statistics, each type of quark appearing in three unitarily equivalent states:

$$q = (q_1, q_2, q_3),$$

which differ in the values of a new quantum number called afterwards the *color*. Since only three quarks were known when the new quantum number was introduced, i.e. u, d, and s, the colored quark model became known at the three-triplet model.

The wave function of an observed baryonic family, which can be approximated for spin unitary symmetry by a fully symmetric 56 component tensor Φ_{abc}, was assumed to be fully symmetric with respect to the color variables of the three constituent quarks:

$$\psi_{ABC}(x_1, x_2, x_3) = \frac{1}{\sqrt{6}} \varepsilon_{\alpha\beta\gamma}\Phi_{abc}(x_1, x_2, x_3), \tag{3.1}$$

where $A = (\alpha, a)$, $B = (\beta, b)$, and $C = (\gamma, c)$.

From this assumption it can be concluded that the observed mesons and baryons are neutral relative to the new quantum number and conform to singlet states, if expressed in terms of unitary $SU^c(3)$ symmetry corresponding to the quantum number. So, for example, the known mesons and baryons are built up from quarks and antiquarks in the following fashion:

$$\begin{array}{ll} \bar{q}^{\alpha}(1)\, q_{\alpha}(2) & \text{mesons,} \\ \varepsilon^{\alpha\beta\gamma} q_{\alpha}(1)\, q_{\beta}(2)\, q_{\gamma}(3) & \text{baryons,} \end{array} \tag{3.2}$$

where the color indices α, β, γ each have only three possible values.

Thus, starting from the colored quark hypothesis, the Pauli principle requirements can be met both for spin quarks as well as hadron spectroscopy.

Colored quarks can have both fractional and integral charges. In the latter case the quarks could be created in high energy particle collisions without violating the basic law of electric charge conservation and would be, generally speaking, unstable particles decaying into known hadrons and leptons [30].

Below we shall discuss the consequences which result from the hypothesis that charge on the quarks is integral (see Sec. 3.7).

Let us stress that the introduction of a new quantum number led to the quarks being seen as normal physical objects accessible to direct or indirect observation.

3.2.2 Dynamic Quark Models

The introduction of colored fermi quarks as fundamental physical particles has opened a way to the dynamic description of elementary particles.

It is the absence of quarks in the free state that has been difficult to explain. Yet explaining this phenomenon, known as quark confinement or non-emission, is one of the crucial problems now facing elementary particle physics. Although the confinement problem will obviously be solved after experiment, a number of attempts have been made to explain the "permanent detention" of quarks inside hadrons in a logically consistent way. In particular, the quark "bag" model has been suggested.

The quark bag model originated in the work done at Dubna (the "Dubna quark bag") [4, 13] and at MIT [14].

A dynamic quark model, whose development started in 1964 at Dubna, was based on the assumption that quarks were very heavy objects bound in hadrons by immense forces. On the one hand, these forces ensure a large mass defect of the quarks in the hadrons, and on the other prevent them from being emitted [1].

---

[1] The description in terms of heavy, composite [11] quarks is, in a sense, complementary to the description in terms of light or "current" quarks

Quark confinement is not unconditional and in principle quarks can be freed if the hadrons acquire a sufficiently high energy.

The dynamic composite model offers a systematic description both of the static observable properties of elementary particles (μ, g_A/g_V, etc.) and the hadron form factors. Note that among the others it gave the first satisfactory explanation of the enhancement of the magnetic moment of a heavy quark bound inside a hadron. This effect could be simply described by the Dirac equation for a quark bound by a scalar field described by a rectangular potential well with $U(r) = -U_0\theta(r_0 - r)$, and in the presence of an external magnetic field **H** by

$$[E + i\boldsymbol{\alpha}(\nabla + ie\mathbf{A})]\psi = \beta M^*\psi; \quad r \leqslant r_0, \tag{3.3}$$

where $M^* = M - U_0$ and $\mathbf{H} = \operatorname{curl} \mathbf{A}$.

Solving equation (3.3) in the limit of an infinitely heavy mass M of the free quark at a fixed value of the effective mass M^*, which hereafter we will assume is zero, we can obtain the following expression for magnetic moment of the bound quark:

$$\mu = \frac{e}{2E} \cdot \frac{4Er_0 - 3}{6(Er_0 - 1)} \simeq 0.83 \frac{e}{2E}; \quad \left(E \simeq \frac{2.04}{r_0}\right). \tag{3.4}$$

We emphasize that the finiteness of the magnetic moment of an infinitely heavy, bound quark can be inferred from an assumption about the scalar nature of the binding potential but does not obtain in the vector case, for example.

This result makes it possible to obtain a good qualitative estimate of the absolute values of the nucleon's magnetic moment, taking one third of the nucleon's mass as the energy of the bound quark and using $SU(6)$ symmetry, i.e.

$$\mu_p \simeq 3 \text{ n.m}, \quad \mu_n \simeq -2 \text{ n.m}$$

and a number of other relations.

Comparing these results with the experimental values reveals the importance of studying the relativistic corrections for the matrix elements of the electromagnetic and weak currents of composite particles. The nature of these corrections can be demonstrated in the most simple and elegant fashion in terms of a model of quasi-independent quarks.

In this model the quarks that make up a hadron move independently within some self-consistent scalar potential $U(r)$, their binding to which compensates for their mass[2]. If the weak and electromagnetic

---

which is a more adequate notion when analyzing a hadron's point structure. Light quark confinement is associated with the non-Abelian nature of the fundamental QCD interaction.

[2] A. Salam has figuratively called this effect an "Archimedian bath" [33].

interactions are introduced in a minimal way [32],

$$i\partial_\mu \to i\partial_\mu + \begin{cases} eA_\mu \text{ electromagnetic interaction,} \\ \dfrac{G}{\sqrt{2}}\tau^\pm\gamma_5 l_\mu^\pm \text{ weak interaction,} \end{cases} \tag{3.5}$$

where A_μ is the electromagnetic potential, $l_\mu^\pm$ charged weak lepton currents, and G the Fermi constant for weak interactions, we will obtain the following for the ratio of the axial to the vector constants of the weak interaction, g_A/g_V, and for the magnetic moment of proton, say,

$$g_A/g_V = -\frac{5}{3}\langle\uparrow|\sigma_z|\uparrow\rangle, \tag{3.6a}$$

$$\mu_p = \frac{e}{2E_q}\langle\uparrow|\sigma_z + L_z|\uparrow|\rangle\,. \tag{3.6b}$$

Here $\sigma_z/2$, L_z, and E_q are respectively the spin and orbital moments and the energy of an individual, bound quark in a nucleon whose projection of total angular momentum is

$$\langle\uparrow|J_z|\uparrow\rangle = \left\langle\uparrow\left|\frac{1}{2}\sigma_z + L_z\right|\uparrow\right\rangle = \frac{1}{2}. \tag{3.7}$$

Whence it is not difficult to find out that

$$g_A/g_V = -\frac{5}{3}(1-2\delta), \tag{3.8a}$$

$$\mu_p = \frac{e}{2E_q}(1-\delta), \tag{3.8b}$$

where the parameter δ,

$$\delta = \langle\uparrow|L_z|\uparrow\rangle = -i\int d^3r\,\psi^*\,[\mathbf{r}\times\nabla]_z\,\psi, \tag{3.9}$$

specifies the value of the relativistic corrections. For an ultrarelativistic case, when $\langle\mathbf{q}^2\rangle/E_q^2 \sim 1$, we have $\delta \sim 1/6$ which yields a correction of the order of 30% for the g_A/g_V ratio. This example shows up the degree to which the effect of relativistic corrections could be essential for predictions of non-relativistic quark models. So-called configuration mixing is another indication of the role of relativistic effects. This is when, in the most general vector representation of a baryon (composed of three quarks), P- and D-waves add to the S-wave contribution [34]. The spin unitary part of the wave function will then not be fully symmetric, being a superposition of the contributions corresponding to 20-, 56-, and 70-tuple representations of the SU (6) group.

The dynamic composite models led to a qualitative explanation and quantitative description of an entire set of transmutational

particle and resonance processes. We specially mention the quark model of electromagnetic and weak meson decays [35, 36] which has been developed using the dynamic approach.

This model offers an explanation of the weak leptonic decays of the pseudoscalar π- and K-mesons and the electromagnetic decays of the vector meson resonances (ρ^0, ω^0, and ϕ^0) into electron-positron pairs as the annihilation of quarks and antiquarks bound inside these mesons. The widths of corresponding decays are governed by the values of the wave functions of bound quark-antiquark pairs in matching coordinates [35], i.e.

$$\Gamma(\pi \to \mu\nu) = \frac{G^2 \cos^2\theta}{2\pi^2} m_\mu^2 \left(1 - \frac{m_\mu^2}{m_\pi^2}\right) |\psi_\pi(0)|^2, \tag{3.10}$$

$$\Gamma(V^0 \to e^+e^-) = \frac{16\pi\alpha^2}{3m_V^2} g_V^2 |\psi_V(0)|^2; \; (V^0 = \rho^0, \omega^0, \phi^0, \ldots), \tag{3.11}$$

where

$$g_\rho = 1/\sqrt{2}; \quad g_\omega = 1/3\sqrt{2}; \quad g_\phi = -1/3;$$

G is the weak Fermi constant, and θ the Cabbibo angle[3].

An analysis of the data on the widths on the meson resonances based on these formulas led to the known conclusion about a dependence of distance scale (effective sizes) on the quantum numbers of a bound system (the Weisskopf-Van Royen paradox), e.g. [36],

$$\frac{|\psi_K(0)|^2}{|\psi_\pi(0)|^2} \simeq \frac{m_K}{m_\pi}, \text{ etc.} \tag{3.12}$$

In the case of the decay $\pi^0 \to 2\gamma$ the annihilation model indicates that the width of this decay is proportional to the number of different quark colors.

The constituent annihilation model and formulas (3.10-3.11), allowing for the first QCD corrections to it form a basis for the contemporary theoretical analysis of the various decay modes of the members of a new family of heavy particles, namely the J/ψ_g-, and γ-mesons.

3.2.3 Hadron Currents and Form Factors

The above examples illustrate the importance of allowing for relativistic corrections when a hadron's local currents and form factors are being constructed. A consistent solution of this problem can only be given using the relativistically covariant description of

---

[3] An allowance for the renormalization of the magnetic moment and axial vector, weak constant of the bound quarks is made by introducing additional factors $(1 - 2\delta)$ and $(1 - \delta)$, into formulas (3.10) and (3.11), respectively.

the system of bound particles within the scope of quantum field theory.

One of the first attempts to construct the local currents of composite particles was performed for [13]. In this work a relativistically covariant equation for mesons and baryons was postulated and this marked the beginning of relativistic, composite quark models.

The equation proposed, the choice of which was dictated by a requirement there should be a generalized $SU(6)$ invariance and that the large mass of the bound quarks was compensated for, enabled many relationships to be obtained for the vertices and form factors of weak and electromagnetic transitions with hadron participation.

It has been proposed in [37, 38] that the three dimensional, dynamic (quasipotential) equations in quantum field theory [39] be used to construct the local currents of the composite particles.

With this approach the matrix current's elements for the bound particle systems will be defined by expressions of the type

$$\langle P', \beta | J_\mu(0) | P, \alpha \rangle = \int \psi^*(\mathbf{q}_i') \tilde{\Gamma}_\mu(P', \mathbf{q}_i' | P, \mathbf{q}_j) \psi(\mathbf{q}_j) \prod_{k=1}^{n} d\mathbf{q}_k' \, d\mathbf{q}_k, \tag{3.13}$$

where $\psi(\mathbf{q}_i)$ is the simultaneous wave function of a bound state in the C.M.S.[4]

The construction of the vertex integral operator $\tilde{\Gamma}(P', \mathbf{q}_i' \mid P, \mathbf{q}_j)$ and the quasipotential equation for the wave function $\psi(\mathbf{q}_i)$ is a basic problem for the simultaneous approach to the description of a system of interacting particles [40-42].

For the case of two spin particles (the quarkonium) the quasipotential wave function equation reduces to the form

$$(\gamma_0 E - \tilde{M}) \psi = 0, \tag{3.14a}$$

$$\gamma_0 \psi \equiv \gamma_0^{(1)} \psi = \gamma_0^{(2)} \psi, \tag{3.14b}$$

where E is the total energy of the system, and $\tilde{M}$ the mass operator which is a function (generally nonlocal) in momentum $\mathbf{q}$, as well as in $i\nabla_q$, $\boldsymbol{\sigma}_{1,2}$, and E [41].

For non-interacting particles the mass operator is merely

$$\tilde{M} = 2W = 2\sqrt{m^2 + \mathbf{q}^2}. \tag{3.15}$$

---

[4] Below in Sec. 3.5 when the asymptotic behavior of hadron form factors are studied for the high momenta transfer range, the three dimensional dynamic equations in the variables of the "light front" (i.e. on the zero plane) are used.

Given an interaction, the operator $\widetilde{M}$ reads

$$\widetilde{M}\psi = 2\sqrt{m^2+\mathbf{q}^2}\,\psi(\mathbf{q}) + \int V(E;\,\mathbf{q},\,\mathbf{q}')\,\psi(\mathbf{q}')\,d\mathbf{q}'. \qquad (3.16)$$

The quasipotential $V(E;\,\mathbf{q},\,\mathbf{q}')$ which enters this expression is specified by the relation

$$\overline{G}^{-1} = \overline{G}_0^{-1} - \frac{1}{2\pi i} V, \qquad (3.17)$$

where $\overline{G}^{-1}$ and $\overline{G}_0^{-1}$ are operators for the inverse bitemporal Green functions of two interacting and two free particles, respectively. In the C.M.S. and in the Foldy-Watthausen representation they are

$$\overline{G}(E;\,\mathbf{q},\,\mathbf{q}') = \Lambda T_{\mathbf{q}}^{+}\left\{\int_{-\infty}^{+\infty} dq_0 \int_{-\infty}^{+\infty} dq_0' G_{\mathbf{p}=0}(E;\,q,\,q')\right\} T_{\mathbf{q}'}^{+}\Lambda. \qquad (3.18)$$

Here

$$T_{\mathbf{q}} = \frac{(m+W-\boldsymbol{\gamma}_1\cdot\mathbf{q})(m+W+\boldsymbol{\gamma}_2\cdot\mathbf{q})}{2W(m+W)}; \quad W = \sqrt{m^2+\mathbf{q}^2} \qquad (3.19$$

is the operator for the unitary Foldy-Watthausen transformation for a two particle system, and

$$\Lambda = \frac{1+\gamma_0(1)\,\gamma_0(2)}{2} \qquad (3.20)$$

is the projection operator of the 16-component bispinors onto a subspace which can have an inverse operator (3.17).

The work [37] demonstrated how the dynamic moments of local vector and axial currents for the bound states of two spin particles can be found from the three-dimensional quasi-potential equations.

Suppose a quark interaction is introduced with the weak homogeneous and time dependent external vector (V_μ^i) and axial-vector (A_μ^i) fields thus:

$$\gamma_\mu\partial_\mu \to \hat{D} = \gamma_\mu\partial_\mu + i\lambda^\alpha V_\mu^\alpha\gamma_\mu + i\lambda^\alpha A_\mu^\alpha\gamma_5\gamma_\mu \qquad (3.21)$$

(here λ^α are the generators of the flavor group).

As a result, the mass operator becomes a function of the external field and can be expanded into the powers of the field, viz.

$$\widetilde{M}_{V,A} = \widetilde{M} + \delta\widetilde{M}_{V,A} + \ldots, \qquad (3.22)$$

where the first non-vanishing correction is controlled by a variation of the two-particle Green function,

$$\delta \widetilde{M}_{V, A} = 2\pi i \overline{G}^{-1} \cdot \delta \overline{G}_{V, A} \overline{G}^{-1} \tag{3.23}$$

and so $\delta G_{V,A}$ can be found by standard perturbational methods.

The vector and axial charges of the particle system,

$$\begin{aligned} Q^{\alpha} &= \left\langle \text{bound} \,\middle|\, \int J_0^{\alpha}(x)\, d\mathbf{x} \,\middle|\, \text{bound} \right\rangle, \\ \mathbf{Q}^{\alpha} &= \left\langle \text{bound} \,\middle|\, \int J_{5i}^{\alpha}(x)\, d\mathbf{x} \,\middle|\, \text{bound} \right\rangle \ (i = 1, 2, 3), \end{aligned} \tag{3.24}$$

can be defined in terms of the total energy variation with the external fields present, i.e.

$$\delta E = V_0^{\alpha} Q^{\alpha} + \mathbf{A}^{\alpha} \mathbf{Q}_5^{\alpha} + \ldots . \tag{3.25}$$

Whence (in the lowest order of perturbation theory for the vertex operator) it follows:

$$Q^{\alpha} = \int \psi^{*}(\mathbf{q}) \{\lambda_1^{\alpha} + \lambda_2^{\alpha}\} \psi(\mathbf{q})\, d\mathbf{q}, \tag{3.26a}$$

$$\mathbf{Q}_5^{\alpha} = \int \psi^{*}(\mathbf{q}) \{\lambda_1^{\alpha} \mathbf{\Delta}_1 + \lambda_2^{\alpha} \mathbf{\Delta}_2\} \psi(\mathbf{q})\, d\mathbf{q}, \tag{3.26b}$$

where

$$\mathbf{\Delta} = \frac{m}{W} \left[\boldsymbol{\sigma} + \mathbf{q} \frac{(\boldsymbol{\sigma} \cdot \mathbf{q})}{m(m+W)} \right], \quad W = \sqrt{m^2 + \mathbf{q}^2}. \tag{3.27}$$

These are exactly the same charges for which the $SU(6)$ algebra was first postulated. In a nonrelativistic approximation, when $\mathbf{q}^2 \ll m^2$, $\mathbf{\Delta} \simeq \boldsymbol{\sigma}$, $SU(6)$ algebra is trivially satisfied for quantities (3.26a, b) given that the two particle states are completely described by the wave functions $\psi(\mathbf{q})$.

However, as was shown above, relativistic corrections may be significant. For example, if the operator $\mathbf{\Delta}$ is averaged over a spherically symmetric state, we find that

$$\langle \mathbf{\Delta} \rangle = \boldsymbol{\sigma} \left(1 - \frac{\langle \mathbf{q}^2 \rangle}{3m^2} \right) + O(1/m^4). \tag{3.28}$$

This agrees with the model of quasi-free quarks, where the renormalization of the axial constant is controlled by a factor in [13], viz.

$$(1 - 2\delta) = 1 - \frac{\langle \mathbf{q}^2 \rangle}{3E_q^2} + O(1/E_q^4). \tag{3.29}$$

In other words, the corrections of the order of $1/m^2$ are needed by relativistic kinematics and are independent of the nature of the particle's interaction.

In the paper by Gell-Mann and Dashen [43] the problem of deriving relativistic quark dynamics from the current algebra in a system with an infinite momentum (the "$P_z = \infty$" system) was posed.

The above analysis allows us to state that the results for the charges on the vector and axial vector currents obtained in these papers can be reproduced without reference to the commutation relations of current algebra at infinite momentum [44].

In fact, the diagonal matrix current elements for the $P_z = \infty$ states are related to the current's diagonal matrix elements at rest by the Lorentz transformation:

$$\langle J_0\rangle_{P_z\to\infty} \to \frac{1}{\sqrt{1-\beta^2}}\,[\langle J_0\rangle_{\mathbf{P}=0}] + \beta\,\langle J_z\rangle_{\mathbf{P}=0}]_{\beta\to 1}. \qquad (3.30)$$

As a result, using (3.26), we find for the limiting values of the vector and axial vector charge matrix elements that

$$\lim_{P_z\to\infty}\frac{1}{P_0}\left\langle\int J_0^\alpha d\mathbf{x}\right\rangle = \int \psi^*(\mathbf{q})\,\{\lambda_1^\alpha+\lambda_2^\alpha\}\,\psi(\mathbf{q})\,d\mathbf{q}, \qquad (3.31a)$$

$$\lim_{P_z\to\infty}\frac{1}{P_0}\left\langle\int J_{50}^\alpha\, d\mathbf{x}\right\rangle = \int \psi^*(\mathbf{q})\,\{\lambda_1^\alpha\Sigma_1+\lambda_2^\alpha\Sigma_2\}\,\psi(\mathbf{q})\,d\mathbf{q}, \qquad (3.31b)$$

where

$$\Sigma_1 = \frac{m}{W+q_z}\left[\sigma_z(1)+\frac{\sigma(1)\cdot\mathbf{q}}{m}\left(1+\frac{q_z}{m+W}\right)\right], \qquad (3.32)$$

$$\Sigma_2 = \frac{m}{W-q_z}\left[\sigma_z(2)-\frac{\sigma(2)\cdot\mathbf{q}}{m}\left(1-\frac{q_z}{m+W}\right)\right].$$

Note that relativistic corrections of the order of $1/m^2$ to axial constant of the two spin particles, which is defined by (3.32), have the same form as before, namely,

$$\langle\Sigma_i\rangle = \sigma_z(i)\left(1-\frac{\langle\mathbf{q}^2\rangle}{6m^2}\right)+O(1/m^4). \qquad (3.33)$$

It is not difficult to show that

$$[\Sigma_1]^2 = [\Sigma_2]^2 = 1, \qquad (3.34)$$

whence the validity of the $SU(3)\times SU(3)$ algebra follows for the charges of vector and axial octet currents of two free quarks. Thus, the three-dimensional dynamic equation can serve as an effective means to check the current algebra of composite particles at infinite momenta.

The results above pertained to the charges, i.e. to the matrix current elements at zero momentum transfer. In [44] it was generalized to slowly varying external fields. This allows the higher dynamic moments of the particle system's currents, such as magnetic and electric dipole moments, to be found.

In particular, evaluating a variation of two spin particles' energy, given homogeneous electric and magnetic fields,

$$\delta E = \mathbf{M}\cdot\mathbf{H} + \mathbf{D}\cdot\mathbf{E}, \qquad (3.35)$$

we find the following expressions for the magnetic and electric dipole moments of the system (in the lowest approximation of perturbation theory for the vertex operators):

$$\mathbf{M} = \int \psi^* (\mathbf{q}) \{e_1\mathbf{m}_1 + e_2\mathbf{m}_2\} \psi (\mathbf{q})\, d\mathbf{q}, \tag{3.36a}$$

$$\mathbf{D} = \int \psi^* (\mathbf{q}) \{e_1\mathbf{d}_1 + e_2\mathbf{d}_2\} \psi (\mathbf{q})\, d\mathbf{q}, \tag{3.36b}$$

where

$$\mathbf{m}_i = \frac{1}{2W} (\mathbf{L} + \boldsymbol{\mu}_i), \quad \mathbf{L} = -i \left[\mathbf{q} \times \frac{\partial}{d\mathbf{q}}\right], \tag{3.37}$$

$$\boldsymbol{\mu}_i = \frac{m+W}{2W} \left[\boldsymbol{\sigma}_i + \mathbf{q} \frac{(\sigma_i \cdot \mathbf{q})}{(m+W)^2}\right], \tag{3.38}$$

$$\mathbf{d}_i = \gamma_0 \left\{\frac{i}{2} \frac{\partial}{d\mathbf{q}} + i \frac{[\mathbf{q} \times \boldsymbol{\sigma}(i)]}{W(m+W)}\right\}.$$

Relativistic corrections to the magnetic moment of a bound quark that were found by averaging over a spherically symmetric state, i.e.

$$\langle \mathbf{m}_i \rangle = \frac{1}{2m} \boldsymbol{\sigma}_i \left(1 - \frac{\langle \mathbf{q}^2 \rangle}{6m^2} + O(1/m^4)\right) \tag{3.39}$$

agree with results from the quasi-free quark model. In this model renormalization of the magnetic moment of the bound quark is governed by the factor

$$(1-\delta) = \left(1 - \frac{\langle \mathbf{q}^2 \rangle}{6E_q^2}\right) + O(1/E_q^4) \tag{3.40}$$

Relativistically covariant quasi-potential equations have been used in [45] to construct form factors for composite particles at arbitrary momentum transfer values.

3.2.4 QCD as the Gauge Theory of Colored Quarks and Gluons

One of the most important implications of introducing color into elementary particle physics is the development of QCD, i.e. the gauge theory of colored quarks and gluons treated as an up-to-date basis of strong interaction theory.

QCD has emerged as a result of bringing the local gauge invariance idea of Yang and Mills [18] into the color $SU^c(3)$ symmetry [17]. An important role in developing QCD was played by the concept of vector particles introduced in 1965 by Nambu. These are the carriers of interaction between the colored quarks and were the prototypes of the QCD vector fields of gluons [16].

Hopes for a theoretical explanation of quark confinement or non-emission[5] are pinned on the non-Abelian nature of a QCD gauge invariance group.

A most important feature of QCD is the asymptotic freedom [46] associated with the discovered weakening of short-range interactions between quarks, i.e. with the increasing momentum transfer:

$$\alpha_s(q^2) \sim \frac{12\pi}{(11N_c - 2N_f)} \frac{1}{\ln q^2/\Lambda^2},$$

where N_c and N_f are the numbers of colors and flavors, respectively, and Λ the fundamental scale.

The QCD's asymptotic freedom corresponds to the idea that the quasi-free quarks, which originate in the dynamic, composite hadron models where light quarks whose mass is "erased" by the interaction, are effectively confined only by the walls of the potential well and remain practically free inside the confinement region (the "Dubna quark bag").

Unifying the heuristic quark bag picture and QCD has led to the development of a number of the modern approaches to hadron dynamics [47, 48].

Theorists at MIT [15] have introduced the assumption that there is some constant density of bag volume energy which is specified by universal parameter $B \simeq \Lambda^4$. This assumption guarantees the stability of a quark bag which has zero complete color with respect to the quark-gluon fields filling the bag.

In a number of works there have been some attempts to substantiate the quark bag model on the basis of a nontrivial topological structure of the QCD vacuum state which is related to instantons (i.e. the effects of tunneling between classically degenerated states which differ by the value of their so-called topological gauge potential number and which conform to zeroth QCD field strengths [47]).

In particular, the application of the finite energy sum rules [49] to analyze the annihilation of e^+e^- pairs into hadrons and to describe the dynamics of quark-antiquark systems (the quarkonium) has led to the following values for the average vacuum fields of light quarks and for a physical vacuum's energy density [50]:

$$\begin{aligned} \langle \bar{u}u \rangle = \langle \bar{d}d \rangle &\simeq -(0.25 \ \text{GeV})^3, \\ \varepsilon = -\frac{9}{32\pi} \langle \alpha_s G^a_{\mu\nu} G^a_{\mu\nu} \rangle &\simeq -(0.25 \ \text{GeV})^4 \end{aligned} \tag{3.41}$$

---

[5] Note that the "elusiveness" of a rigorous proof of color confinement in the non-Abelian theories, in spite of considerable efforts and isolated partial advances in this direction, could indicate that color confinement is, by way of a mathematical analogy, the "Fermat" theorem of contemporary particle theory. In other words, quark confinement or non-emission cannot, generally speaking, be inferred from the first principles of QCD.

(here $G^a_{\mu\nu}$ are the color gluon's field strengths, $a = 1, 2, \ldots, 8$). The negative sign of ε leads to a positive volume density of the bag energy. However, since the quantity B used in describing the hadrons in the quark bag model is almost an order of magnitude less than the energy density of the physical vacuum $|\varepsilon|$ signifies, apparently, that the physical vacuum inside the hadrons is "destroyed" only partially [48].

Note that recently, in connection with studying a possibility of a spontaneous breakdown of color symmetry, the theory is entered by the colored scalar particles of small mass equally with quarks and gluons. A principal corollary of similar theories is the existence of a new hadron family that include the scalar color fields [51-53].

It will be shown below (see Sec. 3.7) that given a condensate of the scalar fields

$$\langle \alpha_s^{1/6} \varphi^+ \varphi \rangle \sim (1 \text{ GeV})^2, \tag{3.42}$$

here $b = \frac{1}{4}\left(11 - \frac{2}{3}N_f - \frac{1}{6}N_s\right)$, N_s being the number of fundamental scalar triplets and allowing for large perturbative corrections in the QCD with scalar fields, the masses of the new hadrons can amount to several tens of a GeV [53]. Observing states such as these would be equivalent to discovering a new flavor, namely, scalar quarks.

3.3 PARASTATISTICS AND COLOR

3.3.1 Parafermi Statistics for Quarks

The first attempt to solve the problem of quark statistics was initiated in 1964 by Greenberg [19] and has been used assuming that quarks are parafermions of rank 3. In this framework, baryons can be explained and described by completely symmetric spin-unitary wave functions. In particular, the baryons built up of three quarks with the identical quantum numbers have been explained, e.g.

$$\Delta^{++}\ (J_z = 3/2) \text{ or } \Omega^-\ (J_z = 3/2).$$

Earlier works (see, e.g. [20]) have emphasized that the use of parafermi statistics for quarks and the introduction of a new quark quantum number (color) together with an appropriate color $SU^c(3)$-symmetry are not equivalent approaches in elementary particle theory and lead, generally speaking, to different physical consequences. Even so, in the last few years a number of works have appeared in which they have been erroneously equated [54].

It is essential to understand this problem properly not only for quark theory, but also for the future development of elementary particle theory. In this section we want to clarify the problem and, using just a brief analysis of the most essential features of both ap-

proaches, we shall show that the use of parastatistics is not equivalent to the introduction of quark color and the appropriate gauge SU^c (3)-symmetry, which form the basis of QCD.

Remember first of all that parafermi statistics are governed by the following set of the trilinear permutation relations

$$[\psi(x), [\bar{\psi}(y), \psi(z)]_-]_- = -2iS(x-y)\psi(z), \tag{3.43}$$

where $S(x-y)$ is the ordinary permutation function of the free fermi fields. As has been shown by Green [55], who was the first to introduce parastatistics, a parafermi field of rank P allows the covariant representation

$$\psi(x) = \sum_{i=1}^{P} \psi_i(x) \tag{3.44}$$

to be made, where the fields $\psi_i(x)$, commonly called the Green parafield components, obey bilinear permutation relations of an anomalous type, viz.

$$[\psi_i(x), \bar{\psi}_j(y)]_\pm = -i\delta_{ij}S(x-y). \tag{3.45}$$

Here the sign "+" (the anticommutator) corresponds to the same components ($i = j$) and, the sign"—" (the commutator) to different ones ($i \neq j$).

Using a parafermi field represented in terms of Green components, it is not difficult to show that in the interesting case of $P = 3$ we have, in contrast to ordinary fermi fields:

$$\psi^2(x) \mid 0\rangle \neq 0, \quad \psi^3(x) \mid 0\rangle \neq 0. \tag{3.46}$$

The last condition solves the problem of baryon spectroscopy that has required it to be possible to place up to the three identical particles into the same quantum state. Note also that the symmetric character of the completely spin-unitary (the SU (6)-part) wave function of observed baryons is because the only composite operator which possesses a unit baryon number in this approach is the operator described in [54, 56] and follows from the ordinary anticommutational Fermi-Dirac relations, namely,

$$[[\psi^a(x), \psi^b(x)]_+, \psi^c(x)]_+, \tag{3.47}$$

where a, b, and c are the spin-unitary field indices.

However, in addition to the ordinary mesons and baryons, for which

$$M \sim [\psi, \bar{\psi}^-]_-, \quad B \sim [[\psi, \psi]_+, \psi]_+, \tag{3.48}$$

a consequence of the normal permutation relations for composite operators leads to systems with the quantum numbers corresponding to a diquark and quark-meson, namely,

$$D \sim [\psi, \psi]_-, \quad F \sim [[\psi, \psi]_+, \overline{\psi}]_+. \tag{3.49}$$

These have no analogues in the spectrum of the hadron states observed.

It should be emphasized that for the parafield to be presented as a sum of Green components does not signify, *per se*, that there are any internal degrees of freedom, i.e. it is only a convenient mathematical means just like the higher spins in the theory of angular momentum can be constructed as a sum of several 1/2 spins. Besides, the anomalous character of the permutation relations of Green components does not allow any physical sense to be directly attached to them.

There is, however, one transformation, i.e. the so-called Klein transformation, that is nonlinear and nonlocal but allows the permutation relations to be reduced to a normal, canonical form. The transformation is of the form [57]

$$\psi_i \to \theta^{-1}\psi_i\theta \equiv \dot{\Psi}_i, \quad \begin{aligned} \Psi_1 &= \psi_i K, \\ \Psi_2 &= i\psi_2 K, \\ \Psi_3 &= \psi_3. \end{aligned} \tag{3.50}$$

The operator K is described by the relations

$$\begin{aligned} \psi_1 K &= K\psi_1; \\ \psi_2 K &= -K\psi_2; \quad KK^+ = K^2 = 1, \\ \psi_3 K &= -K\psi_3; \end{aligned} \tag{3.51}$$

and can be selected in the following form:

$$K = \exp i\pi (N_2 + N_3); \qquad N_i = \int d^3x\, \psi_i^+ \psi_i. \tag{3.52}$$

It can be easily checked that the Ψ_i fields, obtained from the initial Green parafermi field components as a result of the Klein transformation, follow the normal permutation relations:

$$\begin{aligned} [\Psi_i(x), \quad \Psi_j(y)]_+ &= -i\delta_{ij} S(x-y), \\ [\Psi_i(x), \quad \Psi_j(y)]_+ &= 0. \end{aligned} \tag{3.53}$$

Let us stress that the Klein transformation changes the form of the permutation relations and is, therefore, either noncanonical or nonunitary. If the Klein transformation does not change the character of the theory of free fields, then the possibility of applying this transformation to interacting fields generally puts extremely strong restrictions upon the theory, i.e. the so-called superselection rules [58].

In the case we are interested in, i.e. parafermi-statistics of rank 3, as Govorkov showed for the first time in 1966 [59], the relevant superselection rules can be formulated as a requirement of theory invariance given in terms of the normal fermi-fields $\Psi_i = \theta^{-1}\psi_i\theta$ with respect to the transformations of the SO (3)-rotation group of the three-dimensional matter space. In other words, the space of the admissible physical states of the parafields is mapped by the Klein transformation onto the set of those elements Ψ_i of the Fock normal fermi fields space, which are invariant relative to SO (3)-transformations of these fields.

Thus, the observed quantities from the initial theory of parafields of rank 3 will be specified by the vacuum averages of the operators that are invariant relative to SO (3)-transformations of the Ψ_i-fields.

3.3.2 Parastatistics and Gauge Symmetry

If locality or microcausality are required, extremely severe restrictions are imposed on the possible forms for parafield interaction Lagrangians [56]. In particular, attempt to construct a parafermion gauge interaction requires parabosons of the same rank to be introduced [60].

Indeed, out of the two parafermion vector currents,

$$j_\mu(x) = \frac{1}{2}[\bar{\psi}, \gamma_\mu\psi]_-, \tag{3.54a}$$

$$j'_\mu(x) = \frac{1}{2}[\bar{\psi}, \gamma_\mu\psi]_+, \tag{3.54b}$$

only the first is a local operator and can be associated with the electromagnetic current. The second current (3.54b) is generally nonlocal (except for the case when $P = 2$) and satisfies paraboson-type trilinear permutation relations on a spacelike surface.

Let us introduce a paraboson vector field $B_\mu(x)$, which obeys the following permutation relations:

$$\begin{aligned}&[B_\mu(x), [B_\nu(y), B_\lambda(z)]_+]_- \\ &= ig_{\mu\nu}D(x-y)B_\lambda(z) + ig_{\mu\lambda}D(x-z)B_\nu(y),\end{aligned} \tag{3.55}$$

where $D(x-y)$ is the scalar field permutation function. As was the case of the parafermi fields, the field $B_\mu(x)$ can be given as a sum of Green components:

$$B_\mu(x) = \sum_{i=1}^{P'} B_\mu^i(x) \tag{3.56}$$

with the fields $B_\mu^i(x)$ obeying anomalous permutation relations like

$$\begin{aligned}&[B_\mu^i(x), B_\nu^j(y)]_- = -ig_{\mu\nu}D(x-y) \quad (i=j), \\ &[B_\mu^i(x), B_\nu^j(y)]_+ = 0 \quad (i \neq j).\end{aligned} \tag{3.57}$$

Generalizing the Klein transformation ($P = P'$),

$$K = \exp i\pi (N_2 + N_3),$$
$$N_i = \int d^3x\, \psi_i^\dagger \psi_i + \int d^3x\, B_\mu^{-i} \overleftrightarrow{\partial_0} B_\mu^{i\bullet}, \tag{3.58}$$

leads to the normal form of the permutation relations of the Green components both for the parafermi- and parabose-fields.

Following the work in [60], we can see that the most general form of the Lagrangian of a parafermion interaction with vector fields that is compatible with the locality requirement is the following:

$$L = -\frac{1}{4} F_{\mu\nu}^2 - \frac{1}{4} G_{\mu\nu}^2 - \frac{1}{2}\left[\bar{\Psi}, \not{D} \Psi\right]_{-}^{+} e j_\mu A_\mu + g'\left[j'_\mu, B_\mu\right]_+, \tag{3.59}$$

where

$$\not{D} = i\gamma_\mu \partial_\mu - m,$$
$$G_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu + ig\frac{1}{2}\left[B_\mu, B_\nu\right]_-, \tag{3.60}$$

and A_μ and F_μ being the vector potential and strength tensor of the electromagnetic field, respectively.

We stress that the requirement $P = P' = 3$ emerges automatically as a condition of self-consistancy when constructing a local Lagrangian which contains only Yukawa trilinear couplings of the parafermions with the vector fields.

Passing over to Green field components and using the Klein θ-transformation,

$$\Psi_i = \theta^{-1}\psi_i\theta, \qquad B_\mu^i = \theta^{-1} B_\mu^i \theta,$$

from (3.54), and when $g = g'$, we get a Lagrangian of the gauge theory with the SO (3)-symmetry group, thus:

$$L = -\frac{1}{4} F_{\mu\nu}^2 - \frac{1}{4} G_{\mu\nu}^2 - \bar{\Psi} D \Psi + eA_\mu\left[\bar{\Psi}, \gamma_\mu \Psi\right]_-$$
$$+ g\left(\mathcal{B}_\mu, \left[\bar{\Psi} \times \gamma_\mu \Psi\right]_+\right), \tag{3.61}$$

where

$$G_{\mu\nu} = \partial_\mu \mathcal{B}_\nu - \partial_\nu \mathcal{B}_\mu + g\left[\mathcal{B}_\mu \times \mathcal{B}_\nu\right].$$

Note that both fermions (Ψ) and bosons ($\mathbf{B}_\mu$) are transformed here via a three-dimensional (vector) representation of the SO (3)-group.

We shall finish by proving the limited equivalence of the gauge theory with local *SO* (3)-symmetry and the theory of parafields of rank 3 which was given in [60]. As opposed to QCD, this theory has only three gluons and diquarks are absent in the particle spectrum, as are fermions with quark-meson quantum numbers and other exotic hadrons. Besides, the gauge theory with *SO* (3)-symmetry only possesses asymptotic freedom if the number of flavors does not exceed two which contradicts experiment.

Thus, the hypothesis of parafermi statistics for quarks is not equivalent to the introduction of color, or a colored *SO* (3)-symmetry, since it leads to results which are unacceptable on physical grounds.

3.4 HADRON QUARK STRUCTURE AND AUTOMODEL BEHAVIOR AT HIGH ENERGIES

3.4.1 Scale Laws in Particle Physics

The study of interaction processes at high energies and momentum transfers is of primary importance for understanding dynamics of strong interactions and elementary particle structure. The scale properties of these processes have been intensively studied over the last decade, both theoretically and experimentally. All the interaction types, strong, electromagnetic, and weak, exhibit automodel, asymptotic behavior or approximate scale invariance in some form.

In this section we will show how the scale properties can be described starting from the most general, model independent laws of physical similarity, using dimensional analysis, and considering elementary particle quark structure.

Experiments with the high energy particles produced in modern accelerators are now a chief source of information about the elementary particles structure and properties of the fundamental hadron constituents.

As the energy of the interacting particles is increased, which corresponds to ever shorter distances, large numbers of secondaries are produced, since new inelastic channels are opened. The diversity of the processes of interparticle transmutations, the complexity of experimental observation and description of various end products of reactions at sufficiently high collision energies make many of the traditional methods of investigation untenable.

An entirely new approach to studying inelastic high energy particles interactions was proposed by A. A. Logunov in 1967 and has come to take the place of traditional methods. The basis of this method called the inclusive method is that only secondaries of a given sort are observed in the final state of reaction. This allows a model independent description to be made of high energy multiparticle processes based on the fundamentals of quantum field theory [22, 61].

A theory of inclusive reactions has been developed in the works of Logunov and his disciples[6] and has led to the establishment of a number of rigorous asymptotic relations and limitations on the high energy interaction cross sections, and demonstrated the fruitfullness of the approach in analyzing such important regularities as scale invariance.

As soon as the proton accelerator at the High Energy Physics Institute (HEPI) in Serpukhov, which was then the world's largest, was commissioned in 1969, a remarkable feature of scale invariance in the inclusive hadron reactions was discovered when the secondary spectra were studied.

The discovery of scale properties of inclusive processes and their theoretical investigation by R. P. Feynman [62], C. N. Yang [63] and others has deepened our ideas about strong interactions and given a new impetus to the development of the inclusive approach.

The investigations of deep inelastic processes in inclusive electron-nucleon scattering that were carried out in SLAC (Stanford) led to the discovery in 1968 of scale reaction properties (Bjorken scaling). These indicate the existence of a "hard", pointlike structure of a nucleon (the quark-parton structure).

Several years earlier, in 1964, the pointlike behavior of the total cross sections of lepton-hadron interactions was suggested by M. A. Markov [64] from purely theoretical considerations of the dominant role of the channels opening anew, compared to the suppression factor owing to hadron form factors. Experiments on the other big accelerators including those at CERN (Geneva) and Fermilab (Batavia, USA) have supported the pointlike behavior of the deep inelastic scattering of neutrinos and antineutrinos with nucleons. In other words, the effective nucleon size seems to have disappeared in these interactions.

3.4.2 Automodel Principle

In 1969 it was suggested [23] that the experimentally discovered scale properties shown by the electron-nucleon interactions might be assumed to be common for all deep inelastic lepton-hadron processes and could be derived in a model independent fashion using dimensional analysis and physical similarity laws.

An automodel principle was formulated in these works that was universal for describing the scale properties of the widely differing processes of deep inelastic interactions between elementary particles. Essentially, the automodel principle was an assumption that the form factors, and other measurable quantities of deep inelastic processes, were independent at the asymptotic limit of high energies and momentum transfers, of any dimensional parameters (such as

---

[6] See the second paper in the present collection.

particle masses, strong interaction radius, etc.), which can fix the scale of measuring lengths or momenta. Thus, deep inelastic form factors appear to be homogeneous functions of relativistically invariant kinematic variables, whose homogeneity is specified by dimensional analysis[7].

Consider, for example, a deep inelastic process interaction where leptons transfer a momentum q to hadrons having momenta p_i. At the so-called Bjorken limit of $\nu_i \sim s_{ij} \sim |q^2| \gg p_i^2 = m_i^2$ (m_i is mass of the ith hadron) and for fixed values of the dimensionless ratios of the large kinematic invariants ν_i/q^2, s_{ij}/q^2, where $\nu_i = qp_i$, $s_{ij} = p_i p_j$ $(i \neq j)$, an observed parameter, $F(q, p_i)$, of the process under study behaves with a momenta scale transformation

$$q_\mu \to \lambda q_\mu, \quad p_{i\mu} \to \lambda p_{i\mu} \tag{3.62}$$

that corresponds to the automodel principle like a homogeneous function of order $2k$, viz.:

$$F(q, p_i) \to F(\lambda q, \lambda p_i) = \lambda^{2k} F(q, p_i), \tag{3.63}$$

where $2k$ is the physical dimension of the quantity $F(q, p_i)$, i.e.

$$F(q, p_i) = m^{2k}. \tag{3.64}$$

It is not difficult to see that the most general version of the form factor that satisfies these requirements is

$$F(q, p_i) = (q^2)^k f(\nu_i/q_i, s_{ij}/q^2), \tag{3.65}$$

where the function f depends only on the dimensionless ratios of the large kinematic variables that are constant in the Bjorken limit.

For electron-nucleon deep inelastic scattering where the differential cross section is specified by the formula

$$\frac{d^3\sigma}{dk^3} = \frac{4\pi\alpha^2}{q^4}\left[\cos^2\frac{\theta}{2} W_2(q^2, \nu) + \sin^2\frac{\theta}{2} W_1(q^2, \nu)\right], \tag{3.66}$$

(θ is the electron scattering angle in the laboratory frame) the automodel principle leads to scale invariant behavior of the form factors W_1 and W_2. This was found for the first time by Bjorken [65], i.e.

$$\nu W_2(q^2, \nu) = F_2(q^2/\nu), \quad W_1(q^2, \nu) = F_1(q^2/\nu), \tag{3.67}$$

since

$$[W_1(q^2, \nu)] = m^0, \quad [W_2(q^2, \nu)] = m^{-2}.$$

Applying the automodel principle to other lepton-hadron processes has led to a set of important results. In particular, a scale law

---

[7] Automodel behavior in high energy physics is closely analogous to the similarity or automodel property in gas- and hydrodynamics [68] (the term automodel has been borrowed therefrom).

was found which for the first time describes the spectrum of the muon pairs produced in the high energy proton-proton collision $p + p \rightarrow \mu^+ + \mu^- +$ hadrons, namely [67],

$$\frac{d\sigma}{dM} \sim \frac{1}{M^3} \Psi\left(\frac{M}{E}\right), \tag{3.68}$$

where M is the effective mass of the muon pair and E the initial energy of the colliding particles. Experimental studies of this process initiated in 1970 by L. Leberman's group in Brookhaven supported this scale law demonstrating the universality of automodel asymptotic behavior for a wide class of deep inelastic lepton-hadron interactions [68].

Note that in the case of pure, high energy, hadronic collisions the automodel principle leads to so-called Feynman scaling for the inclusive cross sections of the production of secondaries having the limited transverse momenta (with respect to the collision axis), viz.

$$\left(E \cdot \frac{d^3\sigma}{dp^3}\right)_{a+b\rightarrow c+\ldots} = f(p_t, p_z/E). \tag{3.69}$$

That is to say, the collisions in this case result in the inclusive spectra being dependent only on the ratios of the longitudinal momentum components of the isolated secondaries and on the energy p_z/E of primaries. This scale law is derived by analogy with a "flat" explosion in hydrodynamics and uses generalized dimensional analysis of the independent units of measurement of the lengths and momenta along and perpendicular to the collision axis.

Thus, the experimentally observed scale properties of elementary particle interactions can be described in a unified manner based on the automodel principle, which starts from physical similarity laws and dimensional analysis.

At the same time it should be asked to what extent the automodel asymptotic behavior is compatible with the fundamentals and requirements of quantum field theory, such as the conditions of locality, microcausality, and spectrality.

This problem has been fully studied by N. N. Bogolyubov, V. S. Vladimirov, and A. N. Tavkhelidze, who have found sufficient and, in certain cases, necessary conditions for the existence of the automodel asymptote in quantum field theory. One result of this approach has been the establishment of an exact correlation between the automodel asymptote of observables, i.e. amplitudes and cross sections and interaction properties, at extremely short distances [24].

Although providing a theoretical basis from which to understand the general, model independent features of scale regularities, a similar axiomatic approach cannot, quite naturally, pretend to give a concrete form for the functions of the dimensionless ratios of the kinematic variables. These functions characterize the automodel

asymptote and their form is defined by the dynamics of interaction. The additional information required to specify the form of these functions can be found by considering the composite quark nature of hadrons. In particular, the form factors of the deep inelastic lepton-hadron processes in the Bjorken-Feynman quark-parton model are expressed in terms of the distribution functions of the elementary point constituents of hadrons or the partons, i.e. quarks, antiquarks, and gluons.

The quantum mechanical correactions to this model will be discussed below (see Sec. 3.5).

3.4.3 Quark Counting Rules

Some especially interesting and important consequencies have resulted from the idea that hadrons have a composite nature. These consequences are associated with the deep inelastic or exclusive interactions of hadrons, in particular, when binary reactions of the large angle high-energy, hadron scattering are considered. In this kinematic region all the energy and momentum transfers are high and, consequently, we deal with interactions which are concentrated within a region of mainly short distances and time intervals and where the "hard" pointlike quark structure of hadrons must be exhibited in an explicit manner.

In 1973, a general formula was established [25], on the basis of the automodel principle and composite hadron nature. This controls the character of the energy dependence of the differential cross section for an arbitrary binary reaction for large angle high energy ($E = \sqrt{s}$) scattering and the form factors asymptote at high momentum ($Q = \sqrt{-t}$) transfers, viz.

$$\begin{aligned} &\frac{d\sigma}{dt}(ab \to cd) \sim s^{-(n_a+n_b+n_c+n_d-2)}, \\ &F_a(t) \sim t^{-(n_a-1)}, \end{aligned} \tag{3.70}$$

where $n_{i=a,b,c,d}$ are the numbers of elementary constituents of the hadrons participating in the reaction.

This formula is known as a quark counting formula and establishes a direct correlation between the rate of exponential diminishing of the differential cross section for an exclusive binary reaction of large-angle scattering and the energy and degree of complexity of the participating particles, i.e. the number of their elementary constituents.

The discovery of the quark counting formulas has afforded many opportunities to investigate the quark structure of hadrons and light atomic nuclei experimentally [25, 26].

Following [25], we shall dwell briefly on the derivation of formulas (3.70) which were based on dimensional analysis techniques

("dimensional quark counting"). An advantage of this approach to the derivation of quark counting formulas is its universality and independence of special details of a composite structure model of a hadron.

Consider a general binary reaction, $a + b \to c + d$. Assume that at the high energy and momentum transfer limit, particle a behaves like a composite system containing n_a point constituents, quarks, say. The vector of state of such a system can be written thus:

$$| a\rangle = \hat{N}_a \mid n_a, \text{ quarks}\rangle, \tag{3.71}$$

where the symbol $\hat{N}_a$ denotes the operation of multiplying the vector of the free quarks' state by a suitable system wave function and integrating (summing) over the quark variables.

The binary reaction differential cross section v can be given in the form

$$\frac{d\sigma}{dt}(ab \to cd) = \mathrm{Tr}\left(\prod_{i=a,\,b,\,c,\,d} \hat{\rho}_i \frac{d\hat{\sigma}}{dt}\right), \tag{3.72}$$

where

$$\hat{\rho}_i = \hat{N}_i \times \hat{N}_i^{+}, \tag{3.73a}$$

$$\frac{d\hat{\sigma}}{dt} = \frac{1}{s^2} \mid \langle n_a, n_b \mid T \mid n_c, n_d\rangle \mid^2. \tag{3.73b}$$

The dimension of a single particle state that is normalized in a elativistically invariant way is known to be

$$[\mid \text{single-particle}\rangle] = m^{-1},$$

whence the dimensions of the operator factors $\hat{\rho}_i$ and of $d\hat{\sigma}/dt$ which describe corresponding multiquark process follow, i.e.

$$[\hat{\rho}_i] = m^{2(n_a - 1)}, \tag{3.74a}$$

$$\left[\frac{d\hat{\sigma}}{dt}\right] = m^{-2(n_a+n_b+n_c+n_d-2)}. \tag{3.74b}$$

Assuming, in accordance with the automodel principle, that short-range quark interactions are scale invariant, i.e. independent of dimensional dynamic parameters, we arrive at the conclusion of the exponential fall-off of quantity (3.73b) as energy and momentum transfer decrease, as does the differential cross section of the exclusive reaction, i.e.

$$\frac{d\sigma}{dt}(ab \to cd) \to \left(\frac{1}{s}\right)^{n_a+n_b+n_c+n_d-2} f(t/s). \tag{3.75}$$

The function $f(t/s)$ depends only on a ratio of the large kinematic variables, or equivalently on the scattering angle, and is a dimensional quantity with the natural scale being the effective particle size. Thus, the exponential asymptotic law (3.75) indicates a factorization of the effects of short- and long ranges.

The exponential fall off law (3.70) for the hadron form factor can be found by treating a special case of the exclusive reaction, i.e. the scattering of a structureless lepton from a hadron composed on n_a quarks.

Quark counting rules can be generalized for more complicated exclusive reactions as well.

Note that applying the aforementioned considerations to an analysis of inelastic production processes of particles with high transverse momenta, P_t, in high energy hadron collisions leads to pointlike asymptote in the inclusive cross section, i.e. $E\frac{d^3\sigma}{dP^3}\sim P_t^{-4}$ [69-71]. This contradicts available experimental data and has forced a number of authors (see, for example [26]) to search for reasons behind the possible suppression of elementary quark-quark short-range interactions, and also to assume that in these processes, composite systems, i.e. mesons, baryons, etc., behave as elementary constituents.

In particular, if the production of high P_t particles in the inclusive reaction $ab \to c + \ldots$ is conditioned by hard scattering of a quark from a composite system of n_c quarks, then the cross section must have the asymptote $P_t^{-4n_c}$ to match the quark counting rules, i.e. P_t^{-8} ($c =$ pion), P_t^{-12} ($c =$ nucleon), etc.

Further study of these processes, however, shows a deviation from the canonical behavior (P_t^{-4}) of inclusive cross section at high P_t to be correlated with scale invariance (Bjorken scaling) violation in deep inelastic lepton-hadron scattering. The nature of this phenomenon is now the object of intense theoretical studies within the framework of QCD and the composite picture of hadrons.

3.5 QUARK COUNTING RULES AND QCD

3.5.1 Exponential Asymptotic Behavior of Exclusive Processes

In Sec. 3.4 we have shown that the notion of a composite quark structure of hadrons together with the fundamentals of local quantum field theory provides a basis for understanding the major dynamic regularities of high energy particle interactions.

A special position among them is taken by the so-called power laws of particle physics which include the quark counting rules. These establish a direct correlation between the rate of the exponential

fall off of the inclusive reaction cross section with energy and particle complexity, i.e. the number of elementary constituents, namely,

$$\frac{d\sigma}{dt}(ab \to cd) \sim s^{-(n_a+n_b+n_c+n_d-2)}, \tag{3.76}$$

$$F_a(t) \sim t^{-(n_a-1)}. \tag{3.77}$$

A quark counting formula describes, surprisingly well, numerous experimental data on high energy particle scattering and allows immediate information about the number of hadron elementary constituents to be inferred from experiments.

It is of interest to note that experimental results on electron-deuteron scattering at high energies and momentum transfers indicate an applicability of the quark counting fundamentals to the nuclear interactions as well. An analysis of the data of relativistic nuclear physics corroborates this conclusion (see the discussion in Sec. 3.6).

We stress that the exponential asymptotic behavior of exclusive cross sections as predicted by quark counting rules differs qualitatively from Regge mode behavior which results in exponentially low probabilities of high momentum transfer interactions.

The success of quark counting formulas has made it imperative to substantiate it within QCD. Below we shall discuss various approaches to the solution of the problem as well as some new results produced in this direction.

3.5.2 QCD Corrections to Quark Counting Formulas

A number of current works have been dedicated to the corroboration of quark counting rules within QCD. Most of them rest on some summing technique of perturbative QCD diagrams and are applicable, generally speaking, only to short distances (high momentum transfers) [73-76].

The key role in such an approach is played by a statement about factorizing all infrared-divergent contributions (corresponding to long distances) in the form of some structure factors of a hadron wave function type which cannot be found within perturbation theory. The remaining finite factors, which dictate the asymptote of the exclusive amplitudes in the high energy and momentum transfer limit, are given by a renormalized series in a perturbative QCD.

Although a number of difficult questions arise when deriving the factorization and remain inadequately answered, the outcome of this approach to the study of the asymptotics of exclusive processes is worthy of attention. In particular, it is unlikely that it can be rigorously proved that the assumed expansion into a perturbational series and retention of only the diagram's main asymptote in the sum of this series will not destroy the pole structure of bound states of the appropriate Green functions [72].

In Sec. 3.5.3 we shall discuss another approach to the problem of corroborating the quark counting rules which rests on a dynamic treatment of the composite systems in the three-dimensional formulation of quantum field theory on the zero plane.

We now briefly enumerate some of the results of studying exclusive processes in perturbative QCD.

(1) *Meson Form Factors*

The asymptote of the meson electromagnetic form factors have the following form (neglecting higher twist contributions):

$$F(Q^2) = 12\pi C_F \frac{\alpha_s(Q^2)}{Q^2} \times \Big| \sum_{n=0,\,2,\,\dots} a_n \left(\ln\frac{Q^2}{\Lambda^2}\right)^{-\gamma_n}\Big|^2 (1 + O(\alpha_s,\ m/Q)). \qquad (3.78)$$

The logarithmic corrections to the exponential asymptote of the form factor are dictated by the QCD's effective charge behavior, viz.

$$\alpha_s(Q^2) = \frac{g_s^2}{4\pi} \sim \frac{4\pi}{\beta}\left(\ln\frac{Q^2}{\Lambda^2}\right)^{-1}, \quad \beta = 11 - \frac{2}{3}N_f \qquad (3.79)$$

and by the value of the anomalous dimensions

$$\gamma_n = \frac{C_F}{\beta}\left\{1 + 4\sum_{k=2}^{n+1}\frac{1}{k} - \frac{2\delta_{h_q,\,-h_{\bar q}}}{(n+1)(n+2)}\right\} \qquad (3.80)$$

($C_F = \dfrac{n_c^2-1}{2n_c}$), which depend on the way the spins of the quark and antiquark meson constituents are added. If h_q and $h_{\bar q}$ are the corresponding helicities we have

$$\delta_{h_q,\,-h_{\bar q}} = \begin{cases} 0 \text{ for parallel spins } (\rho_{h=\pm 1}), \\ 1 \text{ for antiparallel spins } (\pi,\ \rho_{h=0}). \end{cases} \qquad (3.81)$$

The a_n coefficients in formula (3.78) are defined in terms of the meson's system wave function and cannot, in general, be found using perturbative QCD techniques.

Retaining only the first term in formula (3.78) in the asymptotically high momentum transfer limit, we shall obtain

$$F(Q^2) \to 16\pi f^2 \frac{\alpha_s(Q^2)}{Q^2}\left(\ln\frac{Q^2}{\Lambda^2}\right)^{-\frac{2C_F}{\beta}|h_q+h_{\bar q}|}, \qquad (3.82)$$

where the normalization factor f^2 is controlled by the physically measurable transition constants and it takes the following values for π- and ρ-mesons:

$$f^2 = \begin{cases} f_\pi^2 \simeq (93\ \text{MeV})^2 & \pi \to \mu\tilde{\nu}, \\ f_\rho^2 \simeq (152\ \text{MeV})^2, & \rho \to e^+e^-. \end{cases} \qquad (3.83)$$

According to theoretical estimates, the asymptotic formula (3.82) can be applied far beyond the momentum trasfers attainable in modern accelerators.

We also present the leading asymptote of a pion's transitional form factor that is associated with the $e\pi \to e\gamma$ process. This asymptote has a purely exponential form, viz.

$$F_{\pi\gamma}(Q^2) \to 2f_\pi/Q^2. \tag{3.84}$$

(2) *Baryon Form Factors*

The asymptotic baryon magnetic form factor has the form

$$G_M(Q^2) \to -\langle e_{-\parallel}\rangle \left[C\frac{\alpha_s(Q^2)}{Q^2}\right]^2 \times \left(\ln\frac{Q^2}{\Lambda^2}\right)^{-\frac{2C_F}{\beta}|h_{q_1}+h_{q_2}+h_{q_3}|}, \tag{3.85}$$

where $\langle e_{-\parallel}\rangle$ is the average electric charge of constituent quarks which have helicities h_q that are opposite to the baryonic ones.

Note that at the limit of $SU(6)$-symmetry we get

$$\langle e_{-\parallel}\rangle = \frac{1}{6}(\mu - 3Q) = \begin{cases} 0 & \text{(proton)}, \\ -1/3 & \text{(neutron)}, \end{cases} \tag{3.86}$$

where μ is the magnetic moment and Q the electric baryon charge.

The nucleon electric form factor is exponentially suppressed in contrast to the magnetic one; i.e.

$$G_E/G_M \to \text{const}\cdot\frac{m^2}{Q^2}. \tag{3.87}$$

This is because the amplitudes are suppressed as the helicities change (here $\Delta h = \pm 1$). Baryon form factors which have the helicity $|h| > 1$ are suppressed in an analogous fashion.

(3) *Large Angle Scattering*

The differential cross section of the large angle high energy scattering has the following asymptote:

$$\frac{d\sigma}{dt}(ab \to cd) \to \left(\frac{\alpha_s(Q^2)}{Q^2}\right)^{n-2} \times \left(\ln\frac{Q^2}{\Lambda^2}\right)^{-2\sum_i \gamma_i} f_{ab\to cd}(\theta), \tag{3.88}$$

where

$$n = n_a + n_b + n_c + n_d,$$

$$\gamma_i = |h_i|\, C_F/\beta, \quad i = a,\ b,\ c,\ d.$$

As in every other exclusive process, the amplitudes dominate which satisfy the conservation law of hadron helicities, viz.

$$h_a + h_b = h_c + h_d. \tag{3.89}$$

This law is the result of the vector nature of a QCD interaction.

There is virtually no information about the form of the angular dependence of a binary reaction differential cross sections that is controlled by the functions $f_{ab\to cd}(\theta)$, nor is there any information about the values of the absolute cross section.

The unsolved problem of the so-called "pinch" singularities has to be noted here. Their contributions to large angle scattering amplitudes are characterized by the rate of the exponential decrease of the cross sections, which is smaller than was predicted by quark counting rules (the Landshoff paradox). More investigations must be made before a statement about the suppression of the pinch singularity contribution by Sudakov form factors and the factorization of exclusive binary reactions can be made.

3.5.3 The QCD Description of Composite Systems

We now discuss another approach to the problem of being able to give a consistent account of the quantum mechanical effects when describing high momentum transfer, hadron interactions.

The basis of these methods are the dynamic quasi-potential equations for composite particles given in the "light front" variables which were first introduced by Dirac. The convenience of light front variables is first that when high energy and momentum transfer interactions are studied, particle masses play no significant role and can be taken to vanish[8]. This makes it unacceptable to describe the process in terms of C.M.S. variables. Besides, a theory of composite systems in terms of light front variables most closely approximates the description of interacting hadrons in Feynman's parton model, which imparts the required clearness to rigorous results.

In order to describe the simplest composite system, i.e. a meson which has a 4-momentum P and a set α of quark and antiquark quantum numbers, we choose a gauge invariant two-point function of the Bethe-Salpeter amplitude type [77], i.e.

$$\chi_P(x_1, x_2 \mid C) = \mathrm{Tr}\,\langle 0 \mid T\left(\psi(x_1)\,\bar{\psi}(x_2)\exp\int_{x_1}^{x_2} ig\hat{A}_\mu\,dx_\mu\right) \mid P, \alpha\rangle, \tag{3.90}$$

[8] Obviously, by neglecting exponential corrections of the $O(m/Q)$ type, where Q is the large transferred momentum. This corresponds to the higher twist contributions to the asymptote of the quantum mechanical perturbative theory diagrams.

where the operation T involves both the chronological ordering of the fermion's quark fields and the ordering of the variables of the non-Abelian vector field $\hat{A}_\mu = \frac{1}{2}\hat{\lambda}^a A^a_\mu$ along an arbitrary integration contour C which joins the points x_1 and x_2. In formula (3.90) a trace is taken in the color variables so that all the color of the meson system that is described by the vector of state $|P, \alpha\rangle$ vanishes.

In the case of a baryon system, the gauge invariant amplitude generally depends on three contours C_i $(i = 1, 2, 3)$ which joins the points x_i $(i = 1, 2, 3)$ to some arbitrarily chosen point z, namely,

$$\chi_B = \langle 0 \mid T\,(\mathscr{E}^{abc}(x_1, x_2, x_3 \mid C_i)\,\psi_a(x_1)\,\psi_b(x_2)\,\psi_c(x_3)) \mid B\rangle, \tag{3.91}$$

$$\mathscr{E}^{abc}(x_1, x_2, x_3 \mid C_i) = \varepsilon^{a'b'c'}\left[e^{ig\int_z^{x_1} dx\cdot\hat{A}}\right]^a_{a'}\cdot\left[e^{ig\int_z^{x_2} dx\cdot\hat{A}}\right]^b_{b'}\left[e^{ig\int_z^{x_3} dx\cdot\hat{A}}\right]^c_{c'}, \tag{3.92}$$

with an assumed ordering along each contour.

Note that as a consequence of formulas (3.91) and (3.92) all color of the three quark fields vanishes only at the limit of matching coordinates x_i (or in any other state which is symmetric in coordinates, this corresponding to a requirement of standard SU (6)-symmetry). Yet in general the summed quark color is compensated by gluon field color.

The dependence of (3.90) and (3.91) on the arbitrary contours is obviously an inevitable result of requiring gauge invariance for the QCD description of the composite systems. To justify such an approach we note that gauge invariance must be primarily understood as a requirement of observability. The quantities like (3.90) and (3.91) pertain undoubtedly to the observable, i.e. measurable quantities, since they control the probabilities of such processes as, for example, $\pi \to \mu\bar{\nu}$, $\psi \to e^+e^-$, etc. In addition, they are in the expressions for the exclusive amplitudes of hadron interactions. As will be shown below, we can thus exploit gauge invariant wave functions to construct composite systems' form factors and in doing so factorize the infrared singularities in perturbative QCD in a natural way.

Mandelstam, for example, has used a gauge invariant description for electromagnetic systems [78]. Note, however, that in quantum electrodynamics physical quantities remain invariant if the gauging of either the fermion fields $(\psi \to e^{i\alpha Q}\psi)$ or electromagnetic potentials $(A_\mu \to A_\mu + \partial_\mu\beta)$ alone are changed by virtue of the current conservation of the charged fermion fields, i.e. $\partial_\mu j_\mu = 0$. This means that the formally gauge invariant Green functions can be used for calculations with perturbative QCD.

Since the gluons carry the color charge, only the sum current of fermions and bosons, i.e. $\partial_\mu (j_\mu + ig [A_\nu, F_{\mu\nu}]) = 0$, is known to be conserved in QCD. As a result, those quantities which vary in common gauge transformations of the fermion and color vector fields generally appear to be infrared divergent, i.e. they simply do not exist. The exceptions are those quantities which, by definition, allow only short-range contributions. Thus, the QCD wave function of the bound system whose dynamics is tangibly controlled by long range interactions must be defined in a gauge invariant manner.

Let us discuss briefly how an amplitude like (3.90) depends on the shape of the contour C. By parametrizing the point z_μ on the contour C joining the points $x_{1\mu}$ and $x_{2\mu}$ as

$$z_\mu = z_\mu (s); \quad 0 = s = 1$$

(with $z_\mu (s = 0) = x_{1\mu}$, $z_\mu (s = 1) = x_{2\mu}$) we find that the amplitude variation (3.90) with respect to the contour shape is:

$$\frac{\delta\chi}{\delta C_\mu (z)} = \Gamma_\mu (x_1, x_2, z/C)$$
$$= \operatorname{Tr} \langle 0 \mid T \left\{ \psi (x_1) \bar{\psi} (x_2)\, ig \dot{z}_\nu (s) F_{\mu\nu} (z)\, e^{ig \int_{x_1}^{x_2} dx \cdot \hat{A}} \right\}$$
$$\times \mid P, \alpha\rangle, \quad \dot{z}_\mu (z) = dz_\mu (s)/ds. \quad (3.93)$$

It is not difficult to see that the first variation of quantity (3.90) with respect to the contour's shape is proportional to the amplitude of the bound quark, antiquark, and gluon system which has the quantum numbers of the initial meson system.

Using the QCD equations of motion for the quark-gluon fields we get

$$\begin{aligned} &\gamma_\mu (\partial_\mu + ig\hat{A}_\mu) \psi = 0, \\ &\partial_\nu \hat{F}_{\mu\nu} + ig [\hat{A}_\nu, \hat{F}_{\mu\nu}] = j_\mu, \\ &\operatorname{Tr} \lambda^a j_\mu = g \left(\bar{\psi} \gamma_\mu \frac{1}{2} \lambda^a \psi \right), \end{aligned} \quad (3.94)$$

and we find that the relation

$$\frac{\partial}{\partial z_\mu} \Gamma_\mu (x_1, x_2, z \mid C) = ig \dot{z}_\mu (s)$$
$$\times \operatorname{Tr} \langle 0 \mid T \left\{ \psi (x_1) \bar{\psi} (x_2)\, j_\mu (z)\, e^{ig \int_{x_1}^{x_2} dx\ \hat{A}} \right\} \mid P, \alpha\rangle,$$

links the divergence of Γ_μ with the four quark amplitude.

All the above show that any dependence gauge-invariant ampltudes have on the contours' shape should not be significant when

considering high momentum transfer processes, where the quark counting rules and perturbative computations reveal that any contributions configurations with more than the minimum number of elementary constituents make are suppressed exponentially (the "higher twist" contributions).

The exact equations of motion for the QCD gauge invariant amplitudes of composite systems obviously should define the dependence not only on the end points (the coordinates of the constituent quarks), but also on the shape of the connecting contours. So, for example, a generalization of the Dirac operator on the contour is

$$\hat{D}_x = \gamma_\mu \left[\partial/\partial x_\mu - \int_0^1 ds\, \xi_{\mu\nu}(z)\, \delta/\delta C_\nu(z) \right] \tag{3.96}$$

which is dependent on the following displacement field:

$$\xi_{\mu\nu}(z) = \lim_{\Delta x \to 0} \frac{\Delta z_\mu(s)}{\Delta x_\nu}, \tag{3.97}$$

which is described on the contour C (see Fig. 3.1).

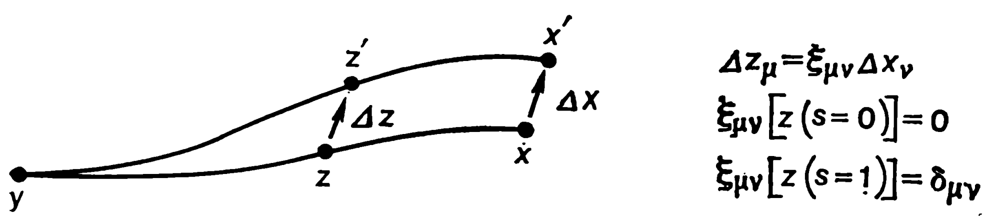


Fig. 3.1. The displacement field on the contour $\xi_{\mu\nu}(z(s))$, which determines the variation of the shape of contour C when the end point is displaced.

The problem of deriving and investigating a whole system of equations and relations which specify a gauge invariant amplitudes lies beyond the limits of the present paper [79].

We shall, however, stress that a dynamic QCD description of a composite system can be developed on the basis of the three-dimensional equations for a wave function related to a gauge invariant amplitude like (3.90), (3.91) for a fixed contour shape.

In particular, when the system's dynamics is described in terms of the light front variables, i.e. on the zero plane,

$$x_\mu = (x_1 - x_2)_\mu, \quad x^+ = x_0 + x_z = 0, \tag{3.98}$$

it is convenient to choose the contour C as lying on hyperplane (3.98). The most convenient gauging here for the vector field is the axial one, viz.

$$A^+ = A_0 + A_z. \tag{3.99}$$

Indeed, the gauge invariant amplitude is readily seen to be explicitly dependent only on the transverse dynamic degrees of freedom of the vector field, i.e.

$$\exp\left[ig\int_C dz\cdot\hat{A}\right]=\exp\left[-ig\int_C d\mathbf{z}_\perp\cdot\mathbf{A}_\perp\right], \tag{3.100}$$

where C lies on the zero plane $(x_1-x_2)^+=0$ and $A^+=0$.

When the interactions have a high momentum, Q, we are interested in the role played by the small impact parameters such that $(x_1-x_2)_\perp\sim 1/Q$. Whence it follows that the exponential factor can be dropped from the definition of the gauge invariant amplitudes like those of (3.90) and (3.91), and the relation will remain accurate to exponential m/Q corrections.

Below we shall briefly show the fundamentals of the theory behind QCD's three-dimensional dynamic equations for composite systems.

If, in the case of meson systems, we use the transformation to the momentum representation, viz.

$$\varkappa_P(q)=\int dx\, e^{ixq}\chi_P\left(\frac{x}{2},\ -\frac{x}{2}\middle|C\right), \tag{3.101}$$

and define the light front variables as

$$x^\pm=\frac{1}{2}(x_0\pm x_z);\quad q^\pm=(q_0\pm q_z);\quad P^\pm=(P_0\pm P_z), \tag{3.102}$$

then we can introduce a quantity

$$\overline{\chi}_P(q^+,\ \mathbf{q}_\perp)=\int_{-\infty}^{+\infty}dq^-\chi_P(q), \tag{3.103}$$

which is dependent on the gauge invariant amplitude (3.90) only in the zero plane, i.e. at $x^+=(x_1-x_2)^+/2=0$.

The meson system wave function will then be defined by the relation

$$\Psi_P(x,\ \mathbf{q}_\perp)=\mathrm{Tr}\,(\hat{\pi}\overline{\chi}_P(q^+,\ \mathbf{q}_\perp), \tag{3.104}$$

where the quantity $x=\frac{1}{2}+q^+/P^+$ varies between $0\leqslant x\leqslant 1$ by virtue of the so-called "projection" properties of χ_P (3.103).

The presence and form of the operator $\hat{\pi}$ in formula (3.104) is controlled by the procedure chosen to project the multipoint Green functions of the Dirac quark fields onto the zero-plane.

A detailed discussion of related problems can be found in [72, 77].

As has been shown in [77], the operator $\widehat{\pi}$ can be chosen, in accordance with the total spin of the quark-antiquark system, in the following form:

$$\hat{\pi} = \begin{cases} \gamma_5\gamma^+ & (h_q + h_{\bar{q}} = 0), \\ \gamma_\perp\gamma^+ & (h_q + h_{\bar{q}} = \pm 1). \end{cases} \tag{3.105}$$

In this way the function (3.104) satisfies the quasi-potential equation thus:

$$\left[M^2 - \frac{\left(\mathbf{q}_\perp + \left(x - \frac{1}{2}\right)\mathbf{P}_\perp\right)^2 + m^2}{x(1-x)} \right] \psi_P(x, \mathbf{q}_\perp)$$
$$= \int_0^1 dx' \int d^2q'_\perp V(x, \mathbf{q}_\perp; x', \mathbf{q}'_\perp)\, \psi_P(x', \mathbf{q}'_\perp), \tag{3.106}$$

where $P^2 = P^+P^- - \mathbf{P}_\perp^2 = M^2$, and the quasi-potential is defined in the standard way using an inversion, i.e. finding the operator which resolves the appropriate four-point Green function when it is projected onto the zero plane of initial and final states.

In the approach to describe composite systems developed here using the gauge invariant amplitudes (3.90), the following expression is natural for the four-point Green functions:

$$G(x_1, x_2 \mid x'_1, x'_2) = \langle 0 \mid T\{0(x_1, x_2 \mid C)\, O(x'_1, x'_2 \mid C'\} \mid 0\rangle, \tag{3.107}$$

where

$$O(x_1, x_2 \mid C = \mathrm{Tr}\left(\psi(x_1)\, \overline{\psi}(x_2)\, e^{ig\int_{x_1}^{x_2} dx\cdot\hat{A}}\right) \tag{3.108}$$

and is the gauge invariant operator which is bilinear in terms of the quark fields and dependent on the arbitrary contour C which joins the points x_1 and x_2. The limiting value of the Green function (5.32) at $g_s = 0$ (in well chosen gauging of the vector fields) will be denoted by $s(x_1, x_2; x'_1, x'_2)$.

According to the general method of the Logunov-Tavkhelidze approach, the relation of the quasi-potential V to the Fourier-images of the Green functions projected onto the zero-plane can in quantum field theory be given in the following symbolic form:

$$(2\pi i)^{-1} V = \overline{G}^{-1} - \overline{S}^{-1} = \overline{S}^{-1}(\overline{SKG})\,\overline{G}^{-1}, \tag{3.109}$$
$$G = S + SKG. \tag{3.110}$$

The projection operation here assumes the axial gauging of the vector fields and the contours C and C' which lie on the zero plane

(see formula (3.10)) as well as the convolution with the operator π prescribed by formula (3.105), namely,

$$\overline{G}(x, \mathbf{q}_\perp; x', \mathbf{q}'_\perp) = \int_{-\infty}^{+\infty} dq^- dq'^- \operatorname{Tr}[\hat{\pi} G_P(q, q') \hat{\pi}]. \quad (3.111)$$

The resolvent of the integral operator (3.111) is defined by the following relation:

$$\int_0^1 dx' \int d\mathbf{q}'_\perp \overline{G}(x, \mathbf{q}_\perp; x', \mathbf{q}'_\perp) \overline{G}^{-1}(x', \mathbf{q}'_\perp; x'', \mathbf{q}''_\perp) = \delta(x - x'') \delta(\mathbf{q}_\perp - \mathbf{q}''_\perp). \quad (3.112)$$

In [74, 77] expression for the contribution of the single gluon exchange diagram to the quark-antiquark interaction quasi-potential has been found by standard perturbational means, viz.

$$\begin{aligned} V &= \overline{S}^{-1} \cdot (\overline{S K_1 S}) \cdot \overline{S}^{-1} + O(\alpha_s^2), \\ K_1 &= \frac{\pi \alpha_s}{K^2 + i0} C_F \gamma_\mu^{(1)} \cdot \gamma_\mu^{(2)} \left[g_{\mu\nu} - \frac{n_\mu k_\nu + n_\nu k_\mu}{(nk)} \right] \end{aligned} \quad (3.113)$$

in the axial gauging, i.e. at $n_\mu^2 = 0$, $\mathbf{n} \cdot \mathbf{k} = k^+ = k_0 + k_z$.

Depending on quark and antiquark helicity values, the single-gluon exchange quasi-potential will be of the form

$$\begin{aligned} &\left\{ \begin{matrix} h_q + h_{\bar{q}} = 0 \\ h_q + h_{\bar{q}} = \pm 1 \end{matrix} \right\} : V(x, \mathbf{q}_\perp; x', \mathbf{q}'_\perp) \\ &= \frac{8\pi G \alpha_s}{x(1-x) y(1-y)} \left[(y(1-y) \mathbf{q}_\perp^2 + x(1-x) \mathbf{q}_\perp'^2) \right. \\ &\times \left(\left\{ \begin{matrix} 1 \\ 0 \end{matrix} \right\} + \frac{\varepsilon(x-y)}{(x-y)} \right) + \mathbf{q}_\perp \cdot \mathbf{q}'_\perp \left. \left\{ \begin{matrix} xy + (1-x)(1-y) \\ 0 \end{matrix} \right\} \right] \\ &\times \left[\frac{\theta(y-x)}{(x-y)\left(\frac{\mathbf{q}_\perp^2}{x} + \frac{\mathbf{q}_\perp'^2}{1-y} \right) - (\mathbf{q}_\perp - \mathbf{q}'_\perp)^2 - i0} \right. \\ &\quad \left. + \frac{\theta(x-y)}{(y-x)\left(\frac{\mathbf{q}_\perp^2}{1-x} + \frac{\mathbf{q}_\perp'^2}{y} \right) - (\mathbf{q}_\perp - \mathbf{q}'_\perp)^2 + i0} \right]. \end{aligned} \quad (3.114)$$

Note that using axial gauging leads to the occurrence of singular expressions like $\varepsilon(x - y)/(x - y)$ in equation (3.114), which should be regularized. As was shown in [74, 77] the Ward identities enable the regularization of the quasi-potential singularities to be related to corresponding Green function renormalizations whilst allowing for their QCD multiloop corrections.

A study of the asymptote of the quasi-potential equations for the high transverse momenta range led Brodsky and Lepage to formulate the so-called evolution equation:

$$\frac{\partial}{\partial \tau} \phi(x, \tau) = \frac{C_F}{\beta} \int_0^1 dy\, V(x, y)\, \phi(y, \tau) \tag{3.115}$$

for the function

$$\phi(x, \tau) \equiv \int_0^{Q^2} d\mathbf{q}_\perp^2\, \psi(x, \mathbf{q}_\perp^2),$$

where

$$\tau = \frac{1}{4\pi} \beta \int_{Q_0^2}^{Q^2} dk_\perp^2\, \alpha_s(k_\perp^2)/k_\perp^2 \simeq \ln\left[\frac{\alpha_s(Q_0)^2}{\alpha_s(Q^2)}\right].$$

The integral kernel of equation (3.115) is determined by the asymptote of the single gluon exchange quasi-potential:

$$V(x, \mathbf{q}_\perp; x', \mathbf{q}'_\perp) \to \frac{4\pi\alpha_s}{x(1-x)} (q_\perp^2) \cdot V(x, y) \tag{3.116}$$

$q_\perp \to \infty$,

$q'_\perp$ is fixed,

and it has the following form:

$$V_1(x, y) = 2\, \frac{1-x}{1-y} \left[\begin{Bmatrix} 1 \\ 0 \end{Bmatrix} + \frac{1}{x-y}\right] \theta(x-y)$$
$$+ 2\, \frac{x}{y} \left[\begin{Bmatrix} 1 \\ 0 \end{Bmatrix} - \frac{1}{x-y}\right] \theta(y-x), \tag{3.117}$$

where the standard Gel'fond-Shilov regularization of the singularity is assumed; for example

$$\left(\frac{\theta(x-y)}{x-y},\ \varphi(y)\right) = \int_0^x dy\, \frac{\varphi(y) - \varphi(x)}{x-y}. \tag{3.118}$$

Note that the eigenfunctions of the integral operator (3.117) are Gegenbauer polynomials, i.e. $\psi_n = G_n^{3/2}(2x-1)$, viz.

$$\int_0^1 dy\, V(x, y)\, \psi_n(y) = (1 + \lambda n)\, \psi_n(x), \qquad \lambda_n = -\beta\gamma_n / C_F, \tag{3.119}$$

and the eigenvalues are defined by the dimensions of the nonsinglet quark distributions γ_n which are determined by a QCD single loop approximation (see formula (3.80)).

If we expand the solution of the evolution equations in the eigenfunctions of the operator $V(x, y)$, we find that

$$\phi(x, \tau) = x(1-x) \sum_{n=0,\,2,\,\ldots} c_n \psi_n(x) \left[\frac{\alpha_s(Q^2)}{\alpha_s(Q_0^2)}\right]^{\gamma_n}. \quad (3.120)$$

This result has been used in [73-75] to study the logarithmic corrections to the exponential asymptote of the meson form factors predicted by the quark counting rules.

3.5.4 The Hadron Form Factors at High Momentum Transfers

Using the pion electromagnetic form factors as examples, we can show how the QCD of composite systems enables the amplitude asymptote of exclusive hadron interactions to be investigated in the high momentum transfer range.

In the dynamic approach of describing composite systems, pion form factors are defined by the following general expression:

$$\begin{aligned} &\langle \pi(P') | J_\mu(0) | \pi(P) \rangle = (P'+P)_\mu F_\pi(t) \\ &= \int_0^1 dx\, dx' \int d^2q_\perp\, d^2q'_\perp \overline{\psi}_{P'}(x', \mathbf{q}'_\perp) \\ &\times \mathscr{T}_\mu(P', x', \mathbf{q}'_\perp | P, x, \mathbf{q}_\perp)\, \psi_P(x, \mathbf{q}_\perp). \end{aligned} \quad (3.121)$$

The vertex operator $\mathscr{T}_\mu = \overline{G}^{-1}\overline{R}_\mu \overline{G}^{-1}$ is defined in terms of the following five-point Green function:

$$R_\mu = \langle 0 | T\{0(x_1' x_2' | C')\, J_\mu(z)\, O(x_1, x_2 | C)\} | 0 \rangle, \quad (3.122)$$

where $O(x_1, x_2, | C)$ is the gauge invariant bilinear operator of (3.108) and $J_\mu = \overline{\psi}(z)\, e\gamma_\mu \psi(z)$ the electromagnetic quark current. The projection of the chosen contour C_i shape onto the zero-plane $R \to \overline{R}$ and the corresponding gauge conditions are discussed in [72].

Since the vertex operator $\mathscr{T}_\mu$ (as opposed to the five-point Green function R_μ) has no poles in the bound states whose existence is controlled by long-range interaction, perturbative QCD techniques can be used to find $\mathscr{T}_\mu$. This is one of the advantages of the approach being developed here.

Something else important is that the factorization of infrared singularities (the "factorization" theorem) allows the hadron form factor to be expressed naturally as an integral over the wave functions which accumulate long range contributions with relevant nonperturbational effects. The integral kernel, $\mathscr{T}_\mu$, is defined by a perturbative QCD renormalized series in the constant $\alpha_s(Q^2)$.

The relation of the factorization theorem becomes more transparent once the expression for the pion factor is transformed to

$$F_\pi \propto \int \bar{\psi}_0 [\bar{S}^{-1} \overline{(G\Gamma G)} \bar{S}^{-1}] \psi_0. \tag{3.123}$$

Here the relations used result from the dynamic equation:

$$\begin{aligned} &\bar{G}^{-1} \cdot \psi = \bar{S}^{-1} \cdot \psi_0, \\ &\psi = \psi_0 + \bar{S} \cdot V \psi, \ \bar{G}^{-1} = \bar{S}^{-1} - V, \end{aligned} \tag{3.124}$$

where ψ_0 is the part of the composite system wave function which corresponds to the contribution of only the "soft", i.e. long range component responsible for the quark confinement. The "hard" (short-range) QCD interactions are taken into account by the potential V.

Let us write out the first few terms of the perturbational series and appropriate Feynman diagrams for the vertex operator viz.

$$\tilde{\Gamma} \equiv \bar{S}^{-1} \cdot \overline{(G\Gamma G)} \cdot \bar{S}^{-1}.$$

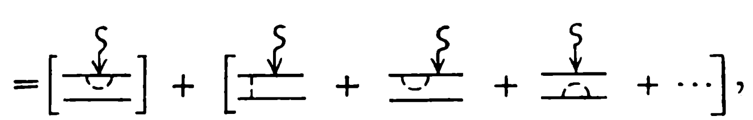
Using the expansions of four-point and reduced five-point Green functions into perturbational series, we get

$$\begin{aligned} G &= S + SK_1S + SK_2S + SK_1SK_1S + \ldots, \\ \Gamma &= \Gamma_0 + \Gamma_1 + \Gamma_2 + \ldots, \end{aligned} \tag{3.125}$$

and find that,

$$\tilde{\Gamma}_0 = \bar{S}^{-1} \left(\overline{S\Gamma_0 S}\right) \bar{S}^{-1} = [\Gamma_0] = [\quad], \tag{3.126a}$$

$$\begin{aligned} \tilde{\Gamma}_1 &= \bar{S}^{-1} \left(\overline{S\Gamma_1 S}\right) \bar{S}^{-1} + \bar{S}^{-1} \left(\overline{SK_1 S\Gamma_0}\right) \bar{S}^{-1} + \bar{S}^{-1} \left(\overline{S\Gamma_0 SK_1 S}\right) \bar{S}^{-1} \\ &= [\Gamma_1] + [K_1 S \Gamma_0 + \Gamma_0 S K_1] \\ &= [\quad] + [\quad + \quad + \quad + \cdots], \end{aligned} \tag{3.126b}$$



where symmetric diagrams are dropped from the last brackets. The brackets in formulas (3.126) signify a transformation to a mass's

surface which assumes, for example, a substitution in "light front" variables like

$$P_i^+ = x_i \cdot P^+, \quad P_i^- = \frac{1}{P^+} \frac{m^2 + \mathbf{p}_{\perp i}^2}{x_i} \tag{3.127}$$

for the momentum components of ith parton in a system having a total momentum P.

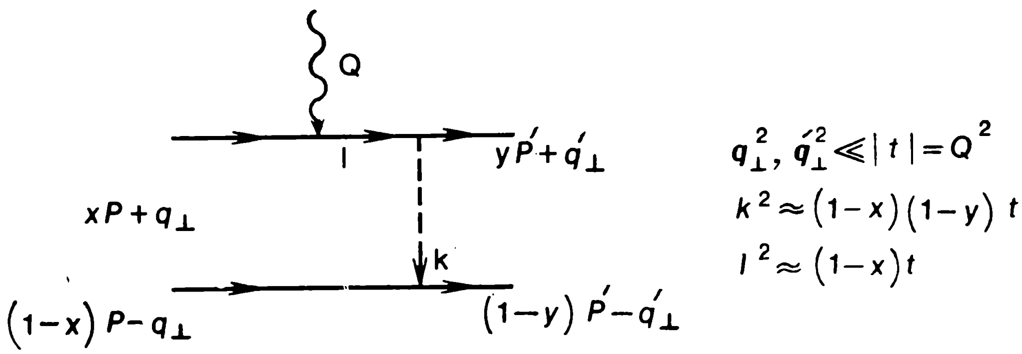


Fig. 3.2. One of the four diagrams of single gluon exchange dominating in the asymptotic limit of high transfer momenta Q.

The main contribution to the asymptote of the pion form factor is given by the diagram of the quark-antiquark single gluon exchange (Fig. 3.2).

Computations yield

$$F_\pi(t) = \frac{4\pi\alpha_s(Q^2)}{Q^2} C_F \left| \int_0^1 \frac{dx}{x(1-x)} \int_0^{Q^2} d^2q_\perp \psi(x, \mathbf{q}_\perp) \right|^2 . \tag{3.128}$$

Assuming further that in the region under consideration, once the wave function $\psi(x, \mathbf{q}_\perp)$ has been integrated over the relative transverse momenta, according to the quark counting rules for the parton distribution, it behaves as follows:

$$\int d^2q_\perp \psi(x, \mathbf{q}_\perp) \simeq C\,[x(1-x)]^{2n_\pi - 3}, \quad x \sim 0.1. \tag{3.129}$$

If $n_\pi = 2$ and we use the condition of normalization to a known constant of decay $\pi \to \mu\tilde{\nu}$, i.e.

$$2\sqrt{n_c} \int_0^1 dx \int d^2q_\perp \psi(x, \mathbf{q}_\perp) = f_\pi, \tag{3.130}$$

we find the leading asymptote of the pion form factor:

$$F_\pi(t) \to \frac{16\pi\alpha_s(Q^2)}{Q^2} f_\pi^2. \tag{3.131}$$

This result coincides with that obtained on the basis of the perturbational techniques (cf. (3.82)). Substituting the evolution equations solution for a partially integrated pion wave function (3.120) into equation (3.128), we find the aforementioned asymptotic expansion of the pion form factor (3.78).

Note that the pion form factor value predicted by formulas (3.82) and (3.131) is several times less than that given by experiment at the momentum transfer values achieved today. This may suggest, first, that the asymptotically high momenta range lies substantially higher than the value customarily assumed, and, secondly that it is necessary to study both the exact form of the composite system wave function and the contributions of the perturbational and non-perturbational corrections to the form factor asymptote.

3.5.5 Angular Dependence of Binary Reactions

The angular dependence of binary reactions contains the important information about the nature of the quark processes that take place in short-range hadron interactions with high momentum transfer.

The study of the angular dependence of the binary reactions in a number of theoretical models has led to the establishment of the so-called generalized quark counting rules [80]. These control the main contributions to the asymptotes of large-angle scattering differential cross sections, in terms of the topologies of the various quark diagrams corresponding to the process:

$$T\,(ab \to cd) \sim \sum_i T_i, \tag{3.132}$$

where i is the quark diagram topology.

Each of the contributions in (3.132) is a homogeneous function (up to the logarithmic corrections) of the large kinematic variables:

$$T_i \sim s^{-\alpha} t^{-\beta} u^{-\gamma}, \tag{3.133}$$

with the exponents, which correspond to the topology of the quark diagram, obeying the condition

$$2\,(\alpha+\beta+\gamma) = \sum_{j=a,\,b,\,c,\,d} (n_j - 1), \tag{3.134}$$

which ensures the appropriate exponential decrease in the differential cross section with increasing energy.

The exponents α, β and γ are individually controlled both by the quark diagram's topology and by the nature of the kinematic singularities of the appropriate helicity amplitudes.

If, as is shown in [80], the kinematic singularities in each reaction channel are isolated in the form

$$T_i(\lambda, \lambda') = \left(\cos\frac{\theta}{2}\right)^{|\lambda+\lambda'|} \left(\sin\frac{\theta}{2}\right)^{|\lambda-\lambda'|} a_i \tag{3.135}$$

(here θ is the scattering angle in the scattering channel, say, $a + b \to c + d$, and $\lambda = \lambda_a - \lambda_b$, $\lambda' = \lambda_c - \lambda_d$, where the λ_i are the hadron helicities) and the generalized quark counting rules are used for the reduced amplitudes a_i, a formula for the angular dependence can be obtained for a wide class of binary reactions.

By way of example, the angular dependence formula for the pion-proton elastic scattering can be given in symbolic form as:

$$\frac{d\sigma}{dt}(\pi p \to \pi p) = \frac{1}{s^2}\left|T_{++}\right|^2 = \frac{1}{s^2}\left|\cos\frac{\theta_s}{2}\underset{A/st^2}{} + \cos\frac{\theta_u}{2}\underset{B/ut^2}{}\right|^2, \tag{3.136}$$

where, by assuming asymptotic γ_5-invariance, a T_{++}-amplitude contribution alone is presented, which conserves the proton helicity. Thus only the quark diagrams are taken into account and they are topologically flat.

Using the expressions for half-angle cosine scattering in the s- and u-channels,

$$\cos\frac{\theta_s}{2} = \sqrt{-\frac{u}{s}}, \quad \cos\frac{\theta_u}{2} = \sqrt{-\frac{s}{u}},$$

we can reduce the angular dependence formula (3.136) to the form $z = \cos\theta_s$)

$$\frac{d\sigma}{dt}(\pi p \to \pi p) = \frac{1}{s^8}\frac{(1+z)}{(1-z)^4}\left|A + \frac{4B}{(1+z)^2}\right|^2. \tag{3.137}$$

We can find the angular dependences for the other binary reactions in a similar way.

The development of QCD computational techniques for exclusive processes means that the problem of corroborating and improving the generalized quark counting rules, which are at present only heuristic, can be approached.

Below we will briefly show, without much substantiation, the results of the angular dependence studies in QCD for the simplest binary reaction, i.e. the elastic quark-pion scattering [81].

The scattering amplitude, assuming γ_5-invariance, has the form

$$T(q\pi \to q\pi) = \left(\bar{u}_{q'}\frac{\hat{p}+\hat{p}'}{2}u_q\right)B(s, t), \tag{3.138}$$

with

$$T_{++} = \sqrt{-us}\,B(s, t), \quad T_{+-} = 0.$$

Treating the pion as a composite system yields

$$B(s, t) = \int_0^1 dx\, dx' \int d^2\mathbf{q}_\perp\, d^2\mathbf{q}'_\perp \bar{\psi}(x', \mathbf{q}'_\perp)\, \mathscr{B}(x', \mathbf{q}'; x, \mathbf{q}_\perp)\, \psi(x, \mathbf{q}_\perp), \tag{3.139}$$

where the integral operator $\mathscr{B}$ is defined in terms of the six-point Green function R_6 projected onto the zero plane, i.e. $\mathscr{B} = \bar{G}^{-1} \cdot \bar{R}_6 \cdot \bar{G}^{-1}$, with $\psi(x, \mathbf{q}_\perp)$ being the pion wave function.

As in the case of the vertex operator $\mathscr{T}_\mu$, which specifies the pion form factor, $\mathscr{B}$ does not contain pole singularities corresponding to the bound states and, as a consequence, can be found using perturbation theory.

The principal asymptotic contributions to the amplitude of the quarkpion scattering will generally be governed by 16 perturbative graphs of order α_s^2. In the limit of many colors, N_c, the contribution of the topologically nonflat quark diagrams appears to be suppressed by the ratio $1/N_c$ compared to the flat ones. In this way we get

$$\mathscr{B} \simeq \frac{(4\pi\alpha_s)^2}{x(1-x)\, y(1-y)} \left\{ C_F^2 \left[\frac{(1-x)(1-y)}{s^2} + \frac{(x+y)}{st} \right] - \frac{1}{2} C_F C_V \left[\frac{1}{st} - \frac{(1-x-y)}{s^2} \right] \pm (s \leftrightarrow u) \right\} \tag{3.140}$$

($\pm$ stand for odd and even amplitude charges).

Integrating (3.140) with the wave functions (3.129), we obtain the following angular dependence for the large-angle quark-pion scattering:

$$\frac{d\sigma}{dt}(q\pi \to q\pi) \simeq \left(\frac{C_F f_\pi \alpha_s}{s} \right)^4 \left(1 - \sec^2 \frac{\theta}{2} \right)^2, \tag{3.141}$$

which fully agrees with the generalized counting quark rules.

3.5.6 A Quark Counting of Anomalous Dimensions of Inclusive Processes

As indicated in the previous section, the assumption about the dominating role of the elementary quark-quark processes and the exact scale invariance of the hadron structure functions leads to point-like asymptote for inclusive cross section at high transverse momenta, P_t.

It was noted, however, that the deviation from the canonical asymptotic law, P_t^{-4}, that has been observed experimentally reflects a scaling violation in deep inelastic lepton-hadron processes.

Studies of the QCD corrections to the point-like, exponential asymptotes of deep inelastic and inclusive reactions at high transverse

momenta have shown that the form of the logarithmic factors which define the deviation from the scaling in these processes is controlled by the hadron's quark structure. The value of the exponent of the large $\log Q^2$ (Q is the large transferred momentum), the so-called anomalous dimensions of the inelastic form factors, turns out to be controlled by the number of quarks that constitute the hadrons that take part in the reactions.

A general formula was established in [82] which defines the leading logarithmic corrections to the canonical point-like asymptote of an arbitrary deep inelastic or inclusive reaction between interacting particles at high momentum transfers in terms of the active and passive quarks which relate to the reacting hadrons.

This formula, called the quark counting rule of anomalous dimensions, expresses an element of the differential cross section of an arbitrary inclusive reaction in the following form:

$$d\sigma = d\sigma_0 \frac{(1-x)^{S-1}}{\Gamma(S)} [(\log Q^2)^{\gamma(S)+r\ln(1-x)}]^H. \tag{3.142}$$

Here $d\sigma_0$ is the appropriate element of the elementary point-like interaction of quarks and/or gluons. and r some group factor equal to 16/25 when $N_c = 3$ and $N_f = 4$. The value of

$$S = \sum_{i(\text{hadrons})} 2(n_i - 1) \tag{3.143}$$

is the double number of the passive quarks (the spectators) related to the participating hadrons and H the full number of active quarks taking part in an elementary hard scattering event, which coincides with the full number of hadrons taking part in the reaction.

The quantity

$$\gamma(n) = -\frac{r}{4}\left[1 + 4\sum_{k=2}^{n} \frac{1}{k} - \frac{2}{n(n+1)}\right] \tag{3.144}$$

is the standard expression (in the single-loop perturbative approximation) for the anomalous dimension of the so-called non-singlet part of the quark-to-hadron fragmentation function.

Consider, by way of an illustration of the general formula (3.142), the deep inelastic electron-nucleon scattering (Fig. 3.3), for which we shall have

$$E\frac{d^3\sigma}{dk^3} \propto \left(\frac{\alpha}{Q^2}\right)^2 \frac{(1-x)^{A-1}}{\Gamma(A)} [(\log Q^2)^{\gamma(A)+r\ln(1-x)}], \tag{3.145}$$

where $Q^2 \sim k_T^2 \gg m^2$, $A \equiv 2(n_A - 1)$, and $\alpha \simeq 1/137$.

For the inclusive hadron jet production with the high transverse momentum $P_t^2 = Q^2$ (see Fig. 3.4) formula (3.142) yields

$$E\,\frac{d^3\sigma}{dP^3} \propto \left(\frac{\alpha_s(P_t^2)}{P_t^2}\right)^2 \frac{(1-x)^{A+B-1}}{\Gamma(A+B)} \times [(\log P_t^2)^{\gamma(A+B)+r\ln(1-x)}]^2, \qquad (3.146)$$

where $A = 2(n_A - 1)$, and $B = 2(n_B - 1)$.

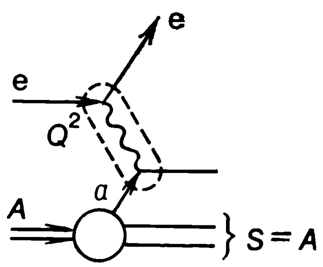


Fig. 3.3. Deep inelastic scattering $e + A \to e + \ldots$, $S = A = 2(n_A - 1)$, $H = 1$.

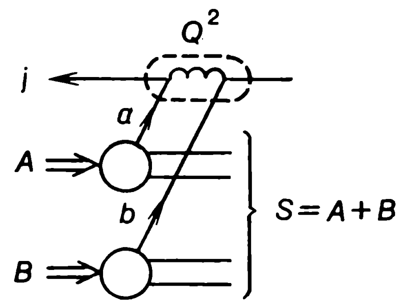


Fig. 3.4. Inclusive formation of the jet $A + B \to j + \ldots$, $S = A + B = 2(n_A + n_B - 2)$, $H = 2$.

An inclusive hadron production with high transverse momentum (Fig. 3.5), as opposed to (3.146), will have the following cross section:

$$E\,\frac{d^3\sigma}{dP^3} \propto \left(\frac{\alpha_s(P_t^2)}{P_t^2}\right)^2 \frac{(1-x)^{A+B+C-1}}{\Gamma(A+B+C)} \times [(\log P_t^2)^{\gamma(A+B+C)+r\ln(1-x)}]^3. \qquad (3.147)$$

This set of the applications of the quark counting rule of anomalous dimensions (3.142) can, if desirable, be continued.

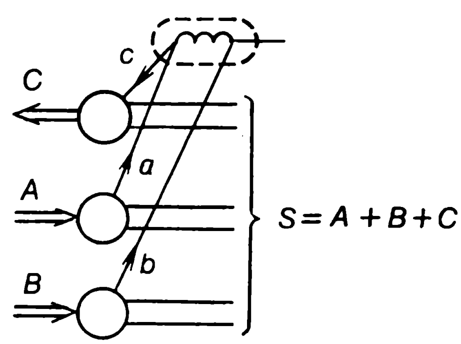


Fig. 3.5. Inclusive hadron production $A + B \to C + \ldots$, $S = A + B + C = 2(n_A + n_B + n_C - 3)$, $H = 3$.

A simple and physically transparent derivation of formula (3.142) can be made starting from the model of the hard scattering of hadron constituents [82, 83].

Indeed, the hard scattering model, for example, in the cases of the inclusive production of a high P_t jet in a hadron collision leads to the following cross section formula:

$$
\begin{aligned}
& E\,\frac{d^3\sigma}{dP^3}\,(A+B\to \text{jet}+\ldots) \\
& =\frac{1}{\pi}\sum_{a,\,b,\,c}\iint dx_a\,dx_b F_{a/A}(x_a,\ Q^2)\,F_{b/B}(x_b,\ Q^2) \\
& \times\frac{d\hat{\sigma}}{d\hat{t}}\,(ab\to cd)\cdot\hat{s}\delta\,(\hat{s}+\hat{t}+\hat{u}),
\end{aligned}
\tag{3.148}
$$

where (see Fig. 3.4)

$$
\hat{s}=(p_a+p_b)^2\sim x_a x_b s;\quad \hat{t}=(p_a-p_c)^2\sim x_a t;\quad \hat{u}=(p_b-p_c)^2\sim x_b u,
$$

and $\frac{d\hat{\sigma}}{d\hat{t}}\,(ab\to cd)$ is the relevant hard scattering cross section defined by the sum of the Born diagrams of perturbative QCD.

The hadron structure function in the first QCD logarithmic approximation (at $x\sim 1$) has the following form:

$$
\begin{aligned}
& F_{a/A}(x,\ Q^2)\propto\frac{(1-x)^{\bar{A}-1}}{\Gamma(\bar{A})}\exp(c\xi), \\
& \bar{A}=A+r\xi=2\,(n_A-1)+r\xi,
\end{aligned}
\tag{3.149}
$$

where

$$
\begin{aligned}
& \xi=\ln\left[\ln\frac{Q^2}{\Lambda^2}\Big/\ln\frac{Q_0^2}{\Lambda^2}\right],\quad r=16/25, \\
& c=r\,(\ln 2-1/2)\simeq 0.12.
\end{aligned}
$$

Substituting (3.149) into (3.148) and changing the integration variables, we obtain

$$
E\,\frac{d^3\sigma}{dp^3}\,(A+B\to\text{jet}+\ldots)\propto\left(\frac{\alpha_s(P_t^2)}{P_t^2}\right)J\,(x_1,\ x_2,\ p_t^2),
\tag{1.150}
$$

where

$$
\begin{aligned}
J & \equiv\frac{(1-x_1)^{\bar{A}-1}\cdot(1-x_2)^{\bar{B}-1}}{\Gamma(\bar{A})\,\Gamma(\bar{B})} \\
& \times\int_0^1\int_0^1\frac{u^{\bar{A}-1}v^{\bar{B}-1}e^{2c\xi}}{[1-u\,(1-x_1)]^2\,[1-v\,(1-x_2)]^2}\,du\,dv \\
& \times(1-x_1x_2)\,\delta\,[uv\,(1-x_1-x_2)-(1-u-v)].
\end{aligned}
\tag{3.151}
$$

Here we used the following notations:

$$x_1 = -t/(u+s) = x_R \frac{\sin^2 \theta/2}{1 - x_R \cos^2 \theta/2},$$
$$x_2 = -u/(t+s) = x_R \frac{\cos^2 \theta/2 \bullet}{1 - x_R \sin^2 \theta/2}, \qquad (3.152)$$
$$x_R = -(u+t)/s = 1 - M^2/s.$$

where M^2 is the square of the missing mass, and θ the scattering angle of the a and b constituents in their center of mass system.

As x_R approaches unity, the integral (3.151) can be computed explicitly as

$$J = \frac{(1-x_R)^{A+B-1}}{\Gamma(A+B)} \frac{\left[(\log Q^2)^{\gamma(A+B)+r \ln(1-x_R)}\right]^2}{\left(\cos^2 \frac{\theta}{2}\right)^{\bar{A}} \cdot \left(\sin^2 \frac{\theta}{2}\right)^{\bar{B}}}, \qquad (3.153)$$

where

$$\gamma(n) = -r\psi(n+1) + c \simeq -\frac{r}{4}\left[1 - \frac{2}{n(n+1)} + 4 \sum_{k=2}^{n} \frac{1}{k}\right],$$
$$c = r\left(\ln 2 - \frac{1}{2}\right) \simeq 0.12$$

and $\psi(n)$ is the digamma function.

It is interesting to note that the dependence of initial hadron momenta on the fraction x_R at $x_R \sim 1$ and the high momenta transfer Q, which follows from formula (3.153), exhibits an effective "collectivization" of the constituents pertaining to different hadrons. Indeed, the whole x_R-dependence enters via the effective structure function of the unified dihadron system $D = (A + B)$ which unifies the A and B constituents, i.e.

$$F_{(a+b)/(A+B)}(x_R, Q^2) \propto \frac{(1-x_R)^{s-1}}{\Gamma(\bar{s})}; \qquad (3.154)$$
$$\bar{S} = \bar{A} + \bar{B} = 2(n_A - 1) + 2(n_B - 1) + 2r\xi.$$

The a and b components are isolated from here and carry away the x_R fraction of the total momentum of the dihadron D.

The contribution of the two-loop corrections and the double logarithmic factors (the quark form factors), which modify formula (3.142) of the hard scattering model near the phase space boundary, was investigated in [83].

3.6 QUARKS IN NUCLEI

3.6.1 Quark Degrees of Freedom of Nuclei

The role of the quark degrees of freedom in describing nuclear phenomena, especially those at high energies and momenta transfer, is a topical item of contemporary nuclear physics.

The starting point of the majority of the works in this direction is an attempt to answer the following question: does the simple fact that the nucleons, which make up known atomic nuclei, are composed of quarks allow us to understand better regularities of purely nuclear phenomena such as the high excitation behavior of nuclear matter, short-range nuclear structure, etc.?

In other words, can atomic nuclei be understood simply as multiquark systems? Under what conditions could quark degrees of freedom be exhibited explicitly?

These questions introduce us to a new and promising branch of research which could lead to radical changes in our notions of the world of atomic nuclei. Undoubtedly, the notion of colored quarks and gluons as the fundamental constituents of matter could shed new light on the nuclear properties of matter and on the nature of nuclear forces.

We shall now briefly discuss only some aspects of the quarks-in-nuclei problem referring to details to the original and review papers.

3.6.2 Nuclear Form Factors at High Momentum Transfer

The most straightforward indication of the quark structure of nuclei stems from the experimental data on evolution of deuteron's electromagnetic form factor at high momentum transfer. This is in a good agreement with the quark counting predictions [27, 28], i.e.

$$F_A(t) \sim t^{-(3A-1)}. \tag{3.155}$$

Here $3A$ is the total number of the valence quarks inside a nucleus composed of A nucleons. Formula (3.155) predicts the following exponential fall in the nuclear form factors, viz.

$$\begin{aligned} &F_{\mathrm{D}}(t) \sim t^{-5},\\ &F_{^3\mathrm{He}}(t) \sim F_{\mathrm{H}^3}(t) \sim t^{-8},\\ &F_{^4\mathrm{He}}(t) \sim t^{-11}, \text{ etc.} \end{aligned} \tag{3.156}$$

Available experimental data enable us to trace a "leveling off" trend of the deuteron form factor $(q^2)^5 F_{\mathrm{D}}(q^2)$ which is normalized in an appropriate way.

The data under discussion have been obtained at SLAC (Stanford, USA) and they concern the effective form factor, $F_{\mathrm{D}}(t) = [A(t)]^{1/2}$, which is controlled by an elastic electron-deuteron scattering cross section [27]:

$$\frac{d\sigma}{dt} = \left(\frac{d\sigma}{dt}\right)_{\mathrm{Mott}} \left[A(q^2) + B(q^2)\tan^2\frac{\theta}{2}\right], \tag{3.157}$$

where ($\eta = q^2/4M_D^2$)

$$A(q^2) = G_0^2(q^2) + \frac{2}{3}\eta G_1^2(q^2) + \frac{8}{9}\eta^2 G_2^2(q^2) \tag{3.158}$$

is the combination of squares of the electric dipole and magnetic quadrupole form factors, respectively.

As the SLAC data show, the deuteron electromagnetic form factor in the $0.8 \leqslant q^2 \leqslant 6.0$ (GeV) transferred momenta range behaves almost exponentially close to $F_D(q^2) = (q^2)^{-5.0\pm0.5}$. This is in remarkable agreement with the quark counting rules and accords with an expected six-quark deuteron structure.[9] Note that for the nuclei, ^{3}He and ^{4}He, the momentum transfers are too sufficiently high as yet to come to any definite conclusions about their quark structures.

3.6.3 Deuteron as a Six-Quark System

The results above show that deuteron's short-range behavior, i.e. at high momentum transfers, is more adequately described in terms of quarks than nucleons.

Does it follow therefrom that deuteron should be considered as a six-quark system, bound together by color QCD forces like, say, a quark bag?

To anticipate a little, note that the analysis of this question leads to the concept "hidden color", a notion which plays a significant role in the physical interpretation of exotic multiquark systems (dibaryons, tribaryons, tetrabaryons, etc.) and in the description of the short-range features of nuclear matter.

Suppose we discuss the properties of a six-quark system as having deuteron quantum numbers and described by the quark bag model. All six quarks are assumed to belong to a ground state with the energy E_q ($j^P = 1/2^+$) $= 2.04/R$ ($m_q \simeq 0$) in a static, spherically symmetric cavity of radius R.

The structure of multiquark system's wave function is defined by the following basic principles:

(1) **the Pauli principle,**

i.e. a complete antisymmetrization of all quark variables including spin, isospin, and color;

(2) **the zero color request,**

which by virtue of an assumed colorlessness of a multiquark system allows only singlet group representations of the colored $SU^c(3)$-symmetry.

For a system of six quarks all of which are at the same energy level, these requirements lead to a wave function which can be written in

---

[9] An analysis of cumulative particle production processes in nuclear collisions at relativistic energies agrees with these conclusions [84].

a symbolic form as

$$\Psi(6q)=\frac{1}{\sqrt{20}}\Big(1-\sum_{\substack{i=1,2,3\\ j=4,5,6}} P_{ij}\Big)\ \overset{\text{colour}}{\begin{array}{|c|c|}\hline 1&4\\\hline 2&5\\\hline 3&6\\\hline\end{array}}\times\overset{\text{spin-isospin}}{\begin{array}{|c|c|c|}\hline 1&2&3\\\hline 4&5&6\\\hline\end{array}}\ , \tag{3.159}$$

where the P_{ij} are the permutation operators of the spin (S), isospin (I), and color quark indices. Moreover, the color and spin-isospin parts of the wave function ($SU^{\mathrm{SI}}(4)$-symmetry) are denoted by the appropriate Young schemes.

Thus, the six-quark system described by the static, spherically symmetric quark bag model has a wave function which corresponds to the 50 component representation of the spin-isospin group of the SU^{SI} (4)-symmetry (the quark analogue of the Wigner symmetry group in nuclear spectroscopy).

The reduction

$$SU^{\mathrm{SI}}(4)\to SU^{\mathrm{S}}(2)\times SU^{\mathrm{I}}(2) \tag{3.160}$$

splits the 50-plet into a sum of $(2J+1,\ 2I+1)$ type terms, i.e.

$$\mathbf{50}=(\mathbf{3},\ \mathbf{1})+(\mathbf{1},\ \mathbf{3})+(\mathbf{5},\ \mathbf{3})+(\mathbf{3},\ \mathbf{5})+(\mathbf{7},\ \mathbf{1})+(\mathbf{1},\ \mathbf{7}). \tag{3.161}$$

The first of these has the quantum numbers of deuteron, the second one, of the virtual $^1S_0$-state, etc.

In connection with the aforementioned experimental indications of deuteron's hard six-quark structure it is natural to ask, "can the six-quark system which has the wave function corresponding to the first term in expression (3.161) be identified with real deuteron?" To answer the question, we must analyze the baryon composition of the six-quark system under consideration. This concerns the relative weights of the various configurations obtained by dividing the initial six-quark system into two subsystems with three quarks each, i.e.

$$[6q]_{B=2}\to[3q]_{B=1}+[3q]_{B=1}. \tag{3.162}$$

In this way we come inevitably across the notion of "hidden color", i.e. the presence in the decomposition of the six-quark system (3.162) of components in which individual three-quark subsystems have nonzero color (colored baryons B_c) [86].

Taking, as a concrete example, a six-quark state having deuteron's quantum numbers, we find the following for the relative weights of the various dibaryon configurations:

$$D_{6q}\to PN\ (1/9);\ \ \Delta\Delta\ (4/45);\ \ B_cB_c\ (4/5). \tag{3.163}$$

Such a large value for the admixture of the hidden color configuration (80%) for the six-quark system, as described by the quark bag model, does not enable it to be identified with real deuteron.

3.6.4 Hidden Color in Nuclear Matter

To reconcile the results of the experiments that measure deuteron's form factor with the large weight of the hidden color component in a six-quark bag, it has to be assumed that real deuteron, which is mainly a weakly bound proton-neutron system, has only low, but finite, probability of going over into the six-quark state as described by the quark bag model. Qualitatively, this can be expressed by the simple formula

$$|D\rangle = a|PN\rangle_{\substack{\text{weakly bound}\\ \text{system}}} + b|6q\rangle, \tag{3.164}$$

where $|a|^2 + |b|^2 = 1$.

The admixture of the six-quark component in deuteron can be determined from the experimental data on deuteron's electromagnetic form factor at high momenta transfer. These yield [85] $\alpha = |b|^2 \simeq 7\cdot 10^{-2}$. Using this value as well as the weights of various components in the six-quark bag with deuteron's quantum numbers, we can readily find an estimate of the exotic admixtures in a real deuteron, namely

$$\begin{aligned} &\Delta\Delta\text{-component} \sim 0.6\%, \\ &\text{“hidden color”} \sim 5.6\%. \end{aligned} \tag{3.165}$$

This estimate obviously has a qualitative character and does not allow for a number of important details, such as the presence of D-waves, the weak non-orthogonality of two vectors of state in formula (3.164), etc.

In any case we conclude that the ordinary, orthodox theory of nuclear matter, which does not allow for the color of fundamental constituents, quarks and gluons, is incomplete.

The search for the experimentally observed consequences of hidden nuclear color is one of the most important problems of relativistic nuclear physics.

One feasible way of research could be to study deuteron electrodisintegration at high energy and momenta transfer. This produces two baryons or hadron jets which are emitted at large angles and have the quantum numbers of the colored three-quark systems contained in the deuteron. Starting from an assumption about the smallness of the ranges corresponding to the dynamic mechanism of the

processes under consideration, we must expect the following relation etween the various reaction channels.

$$d\sigma(\gamma^* D \to PN) : d\sigma(\gamma^* D \to \Delta\Delta) : d\sigma(\gamma^* D \to B_c B_{\bar{c}}) = 5:4:36. \tag{3.166}$$

We refer the reader for details to the original papers and now only show that the differential cross section of the deuteron electro-disintegration in exclusive channels behaves in the following manner in the high momentum transfer (Q) range, if allowance is made for QCD corrections:

$$\frac{d\sigma}{dQ^2} \sim \left(\frac{\alpha_s(Q^2)}{Q^2}\right)^{11} (\log Q^2)^{-\gamma}, \tag{3.167}$$

where

$$\gamma = \frac{2C_F}{\beta}(|h_D| + |h_{B_1}| + |h_{B_2}|)$$

is the anomalous dimension (h_D and h_{B_i} are the helicities of the deuteron and final baryons). The rate of the exponential fall-off in cross section (3.167) as momentum transfer increases agrees with the quark counting rules and consists of the canonical exponent which is equal to two plus $\sum_{i=D,B^1 B_2} (n_i - 1)$.

The search for hidden color in nuclei is difficult both because of the smallness of the relevant effects and because it is not possible directly to observe the hidden color. The discovery of more explicit manifestations of quark color and QCD fields in nuclei would, therefore, be of great importance. We would like to indicate in this connection the interesting possibility that highly excited states of nuclear matter might exist and which could be mainly pure hidden color excitations [29].

The analysis performed on this shows that at excitation energies of the order of $\Delta E \sim 0.5$ GeV, the state widths are controlled by the hidden color "discharge" via a single-gluon exchange and comprise the values $\Gamma_{c\bar{c}} \lesssim 10$ MeV, thus making them accessible for observation.

3.6.5 Quantum Theory of Nuclear Forces

When discussing the role of quark degrees of freedom in the description of purely nuclear interactions, it would be quite natural to ask whether it is possible to understand the principal features of nuclear forces from a knowledge of fundamental QCD interactions. Is it possible, in particular, to determine the nuclear parameters and hadron-hadron interaction constants such as, say, the pion-nucleon coupling constant, in terms of the principal QCD constants?

As we have shown above, the quark model provides a natural explanation of a short-range repulsive core of forces acting between two nucleons. As is indicated in a number of texts, quark exchange between different nucleons is one source of exotic nuclear components and, in addition, it leads to the occurrence of multi-nucleon forces in nuclear matter.

Below we shall give some of the arguments as to how QCD could explain the long-range components of nuclear forces. This is generally done in a phenomenological manner as an exchange of "white" (i.e. colorless) particles, i.e. pions, vector mesons, etc. This is a difficult problem, inasmuch as at long ranges we come across the strong coupling and confinement phenomena.

To simplify the issue, we shall try to answer a simple question, "do two quark bags interact at interbag separation distances exceeding two bag radii?" Obviously this should be understood in a positive way, since it is difficult to imagine a quantum object as having a rigidly fixed boundary. Indeed, the quantum fluctuations of a quark bag's surface over time will result in the formation of the interbag joins (where the color fields are non-vanishing) which would permit quarks to tunnel from one quark bag to another. The tunnelling probability has an exponentially low value at large relative separations, R, between the bag centers, namely, $\sim\exp(-\mu R)$ when $R \gg 2a$ (a is the bag radius), this corresponding to the Yukawa interaction fall-off law.

As is shown in (29], the evolution of a system of two quark bags separated by the distance R ($R > 2a$) is described by the amplitude:

$$\langle R|\overline{e}^{iHT}|R\rangle = \int d^3\sigma \int\limits_{\sigma \text{ is fixed}} d\psi\, d\overline{\psi} dA e^{iS[\psi, A]}. \tag{3.168}$$

Here σ is the three-dimensional hypersurface in space-time that confines the domains with non-zero functions of the colored quark-gluon fields. Over large evolution times, T, amplitude (3.168) has an asymptote at $\sim\exp[-iTU(R)]$, where $U(R)$ is the interaction static potential of the bags. The amplitude of a single tunnelling act of a quark pair from different bags through the joint under the potential barrier $\Delta E \gtrsim 1/\tau$ (τ being the fluctuation time or joint thickness) is proportional to the product $T\cdot U(R)$. According to the quasiclassical nature of the tunnelling, this amplitude is exponentially small, namely,

$$U(R) \sim \exp[-S_0(R)], \tag{3.169}$$

where $S_0(R)$ is governed by the action integral taken over the region of joint and minimized over its thickness at a large and fixed R.

As can be shown [29]

$$S_0(R) \to \mu R \text{ as } R \to \infty, \tag{3.170}$$

where μ is defined by the energy spectrum minimum of the quark-antiquark system, i.e. $\mu = m_\pi$. The result is nothing but the long-range part of the Yukawa interaction which can thus be explained by fundamentals QCD.

3.7 BROKEN COLOR SYMMETRY AND INTEGRAL-CHARGED QUARKS

3.7.1 The Problem of Quark Charges

Even the first work on the three-triplet model have shown that colored quarks could possess both the fractional and integral electric and baryon charges.

The assumption of an exact color symmetry and the non-observability of color is compatible only with fractional charges. However, owing to quark confinement, the straightforward detection of the electric charge of individual quarks is a major problem.

Charge is known to play a double role in quantum electrodynamics. On one hand, it is an integral of motion, whose values govern the states of observed particles, and on the other, it is the interaction constant which is renormalized due to vacuum polarization effects. It is perhaps possible to speak of quark charges within the framework of QCD as only effective constants that characterize the electromagnetic quark interaction at sufficiently small separation distances. Therefore, the values of the electric charge on quarks can be established by comparison with the electric charges on leptons at the same separation distances.

In the case of integrally-charged quarks, color symmetry is not as yet being broken (either locally or globally) at least in electromagnetic interactions. Indeed, in models which have integral-charged quarks the electromagnetic current is the sum of the $SU^c(3)$-group singlet and octet terms, i.e.

$$J_\mu^{\mathrm{em}} = J_\mu^{c=0}\,(1) + J_\mu^{c\neq 0}\,(8). \tag{3.171}$$

So, in the three-triplet model, i.e.

$$\begin{aligned} u &= (u_1,\ u_2,\ u_3), \\ d &= (d_1,\ d_2,\ d_3), \\ s &= (s_1,\ s_2,\ s_3) \end{aligned} \tag{3.172}$$

integral quark charges,

$$Q_u = (1,\ 1,\ 0);\ Q_d = Q_s = (0,\ 0,\ -1), \tag{3.173}$$

are chosen according to the formula

$$Q_q = Q_{\mathrm{GMZ}} + Y_c. \tag{3.174}$$

Here Q_{GMZ} is the standard fractional charge in the Gell-Mann-Zweig model, and

$$Y_{c} = \frac{1}{\sqrt{3}} \lambda_8 = (1/3,\ 1/3,\ -2/3) \tag{3.175}$$

is one of the generators (the 8th component) of the color group.

Generally, we have in integral charged quark model

$$Q_{\mathrm{q}} = Q_{\mathrm{GMZ}} + Q_{\mathrm{c}};\ Q_{\mathrm{c}} = T^{a} \cdot C_{a}, \tag{3.176}$$

where T^a is an array of the eight generators of the $SU^{\mathrm{c}}(3)$-group in the basic representation. It follows from the hermiticity condition of electric charge, i.e. $Q_{\mathrm{c}}^{+} = Q^{\mathrm{c}}$, that if an appropriate selection of the representation basis is made, the color part of the charge Q_{c} can be reduced to the diagonal form viz.

$$Q_{\mathrm{c}} = aT_3 + b\frac{1}{\sqrt{3}} T_8 = \left(a + \frac{b}{3},\ -a + \frac{b}{3},\ -\frac{2b}{3}\right). \tag{3.177}$$

The requirement that the charges be integral leads to $2a$ and $(a + b)$ being integers. Quark charges do not exceed unity at $a = 0$ or $b = 1$ and this corresponding to selection (3.175). The same is valid when $a = b = -1/2$. The last case, however, also converts into (3.175) when the first and third colored quarks are permutated.

It should be stressed that according to the Gell-Mann-Nishijima formula integral-charged quarks must be correlated with integral baryon charges, i.e.

$$Q_{\mathrm{q}} = T_3 + \frac{1}{2} B_{\mathrm{q}};\quad B_{\mathrm{q}} = (1,\ 1,\ -1). \tag{3.178}$$

The dependence of quark charges on their color state obviously leads to a breakdown of global color symmetry in electromagnetic interactions. This is exhibited, for example, in the mass-splitting of colored quark triplets, etc. However, as has been demonstrated in the works cited, hadron neutrality relative to the color (i.e. matching the observed mesons and baryons with singlet color wave functions), guarantees the disappearance of any manifestation of the aforementioned breakdown of global color symmetry for all observed hadron characteristics (charge, magnetic moment, form factor, etc.).

However, the following two important items must be emphasized. First, it appears to be possible in electromagnetic interactions to excite particle states with a nonzero color, having an energy which is assumed to be fairly high, much higher compared to the mass spectrum of the observed mesons and baryons. Secondly, the integral nature of quark's electric and baryon charges makes it possible for them to transform into leptons and other observable particles. We would eventually draw some conclusion about the instability of quarks which could explain the negative results in the search for

them both in the environment and in accelerator experiments. In this way, as Pati and Salam have demonstrated [89], quark instability would not contradict nucleons' high stability or the extreme suppression of the observed effects of baryon charge non-conservation.

In allowing for the QCD quark interaction described by non-Abelian gauge theory, the introduction of integral-charged quarks poses a new problem. The straightforward selection of a quark's electromagnetic current in the form of (3.171) would bring about a breakdown of the local gauge $SU^c(3)$-symmetry and, as a consequence, an inability to normalize the theory. The problem is removed once spontaneously broken color $SU^c(3)$-symmetry is considered. This requires new degrees of freedom to be introduced into the theory, e.g., Higgs color scalar fields.

3.7.2 Spontaneously Broken Color Symmetry

The simplest model of strong and electromagnetic interactions with the spontaneously broken color symmetry and integral-charged quarks is the gauge model based on the $SU^c(3) \times U(1)$-group with the scalar triplet φ_a.

The Lagrangian of the model has the following form [31]:

$$
\begin{aligned}
L = &-\frac{1}{4}(F^0_{\mu\nu})^2 - \frac{1}{4}(F^a_{\mu\nu})^2 + L_\varphi \\
&+ \bar{q}\left(i\hat{\partial} + \frac{g}{\sqrt{3}}\lambda^a\hat{A}^a + g'Y_q\hat{A}^0 - m_q\right)q \\
&+ \bar{l}\,(i\hat{\partial} + g'Y_l\hat{A}^0 - m_l)\,l,
\end{aligned}
\tag{3.179}
$$

where

$$
\begin{aligned}
&F^0_{\mu\nu} = \partial_\mu A^0_\nu - \partial_\nu A^0_\mu; \quad \hat{A}^0 = \gamma_\mu A^0_\mu; \\
&F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + gf^{abc}A^b_\mu A^c_\nu; \quad \hat{A}^a = \gamma_\mu A^2_\mu,
\end{aligned}
\tag{3.180}
$$

and L_φ is the Lagrangian of the Higgs scalar field:

$$
\begin{aligned}
L_\varphi = &\left|\left(\partial_\mu - i\frac{g}{\sqrt{3}}\lambda^a A^a_\mu - ig'Y_\varphi A^0_\mu\right)\varphi\right|^2 \\
&- h\,(\varphi^+\varphi)^2 - m^2\varphi^+\varphi.
\end{aligned}
\tag{3.181}
$$

If the $SU^c(3) \times U(1)$-symmetry is not broken, the A^0_μ field has to be identified with a photon and so the electromagnetic properties of leptons and hadrons result in the following choice of hypercharges:

$$Y_l = Y_\mu = -1; \quad Y_u = 2/3; \quad Y_d = Y_s = -1/3.$$

When $m^2 < 0$, the symmetry is spontaneously broken, the nonzero vacuum averages of the φ_a fields being chosen in the form

$$\langle\varphi_a\rangle\,(0,\ 0,\ v/\sqrt{2}), \tag{3.182}$$

where v equals $(-m^2/n)^{1/2}$ in the tree approximation.

To match this choice, the new vacuum must be invariant with respect to the $SU'(2)\times U'(1)$-group that is composed of elements like

$$g=\exp i\left\{\lambda^a\omega^a+(\lambda^8+2/\sqrt{3})\,\omega\right\},\quad a=1,\ 2,\ 3. \tag{3.183}$$

As a result of a spontaneously broken symmetry, the vector bosons,

$$A_{45}^{\pm}=\frac{1}{\sqrt{2}}(A_4\pm iA_5),\quad A_{67}^{\pm}=\frac{1}{\sqrt{2}}(A_6\pm iA_7), \tag{3.184a}$$

$$U=A^8\cos\theta-A^0\sin\theta,\qquad \tan\theta=\frac{g'}{g}\left(\frac{3Y_\varphi}{2}\right),$$

and the scalar field,

$$\chi=\frac{1}{\sqrt{2}}(\varphi_3+\varphi_3^*-\sqrt{2}\,v), \tag{3.184b}$$

gain mass, viz.

$$m_{45}^2=m_{67}^2=g^2v^2/3,\qquad m_\chi^2=2\,|\,m^2\,|=2hv^2$$
$$m_v^2=\frac{4v^2}{g}\left(g^2+g'^2\left(\frac{3Y_\varphi}{2}\right)^2\right). \tag{3.185}$$

The vector field,

$$A\ =\ A^0\cos\theta\ +\ A^8\sin\theta, \tag{3.186}$$

is the only one massless field to be associated with leptons. For this reason vector field (3.186) must be identified with a photon.

It can be easily derived that the electric charges on quarks and leptons in the case of a spontaneously broken symmetry are equal to

$$Q_e=Q_\mu=e;\quad Q_q=e\left(Y_q+Y_c\left(\frac{3Y_\varphi}{2}\right)\right), \tag{3.187}$$

where $e\ =\ g'\cos\theta$.

If we choose $Y_\varphi\ =\ 2/3$, then all the particles in the theory will have integral charges.

The theory under consideration differs from standard QED in two ways. First, additional scalar degrees of freedom have been introduced which are necessary to break the color symmetry. Second, as a result of the spontaneous breakdown, the lepton symmetries acquire an additional short-range interaction (the U-boson exchange),

and photons become "colored", i.e. they interact in different ways with quarks of different colors.

The Lagrangian mass of the U-boson, m_U, can be chosen as the characteristic scale of the symmetry breakdown. It is known that at momentum transfers when $Q^2 \gg m_U^2$ the initial $SU^c(3) \times U(1)$-symmetry is reestablished and the effective quark charge becomes integral. In the high momentum transfer range the theory differs from QCD only in the occurrence of scalar particles.

The choice of the parameter m_U has been investigated in [31, 53], where the limit $m_U \ll 1$ GeV was established by analyzing the sum rules for the annihilation of e^+e^--pairs into hadrons [49], and by studying the radiative decays of heavy meson resonances and the corrections to the anomalous magnetic moments of muons.

Thus, a model with a spontaneously broken color symmetry and integral-charged quarks does not contradict the available experimental data provided the symmetry breakdown originates at values of momenta beyond the asymptotic freedom range or at distances comparable to the QCD confinement radius.

Lepton electrodynamics simply indicates that the scale of symmetry breakdown cannot be too large. Actually, when $m_U \gtrsim 1$ GeV, a weak coupling mode will act over all distances in strong interactions since the strong interaction constant α_s is "frozen" at the energies lower than a U-boson's mass, i.e. [51]

$$\alpha_s(Q^2) \simeq \alpha_s(m_U^2) \lesssim 1,$$
$$Q^2 < m_U^2.$$

Thus, the U-boson must be observed (with a mass close to the Lagrangian mass) as a resonance in the $e^+e^- \to \mu^+\mu^-$-annihilation and having an electronic width

$$\Gamma_{e^+e^-} \simeq \frac{4\alpha^2}{9\alpha_s(m_U^2)} m_U \gtrsim 70 \text{ keV} \tag{3.188}$$

when

$$m_U = 1 \text{ GeV}, \ \alpha_s(m_U^2) \simeq 0.3.$$

The results of an experimental search of the resonances in the e^+e^--annihilation at the current energies exclude this possibility. In addition, the assumption that m_U values could exceed a few GeV contradicts the achievements of the SU^c (3)-symmetry in classifying hadrons.

The occurrence of scalar charged fields in the models considered brings additional contributions to the asymptote of the total cross section of the annihilation of e^+e^--pairs into hadrons, viz.

$$R = \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)} \to R_q^{ac} + R_s^{ac}, \tag{3.189}$$

where

$$R_q^{ac} = \sum_q Y_q^2 = 2 \quad \text{(for } u\text{-, } d\text{-, and } s\text{-quarks)}, \tag{3.190a}$$

and

$$R_s^{ac} = \frac{1}{4} \sum_\varphi Y_\varphi^2 = \frac{1}{3} \quad \text{(for a scalar triplet)}. \tag{3.190b}$$

It is important to emphasize that $\sqrt{s} \gg m_U$ is an insufficient condition for formulas (3.189) and (3.190) to be applied to the description of experimental data, i.e. it is necessary to require the corrections to be small including both the perturbational (logarithmic) and non-perturbational (exponential) corrections to the leading asymptote of the annihilation cross section.

The allowance made for two-loop and three-loop corrections in the $\overline{\text{MS}}$ renormalization scheme in the model which has a triplet of scalar fields ($Y_\varphi = 2/3$) [87] yields

$$R_q^{ac} = \sum_q Y_q^2 \left\{1 + \frac{\bar{\alpha}_s}{\pi} + 1.47 \left(\frac{\bar{\alpha}_s}{\pi}\right)^2 + \ldots\right\}, \tag{3.191a}$$

$$R_s^{ac} = \frac{1}{3} \left\{1 + \frac{4\bar{\alpha}_s}{\pi} + 36.6 \left(\frac{\bar{\alpha}_s}{\pi}\right)^2 + \ldots\right\}. \tag{3.191b}$$

Here the effective quark-gluon coupling constant $\bar{\alpha}_s$ is defined in the following way:

$$\frac{\bar{\alpha}_s}{\pi} = \frac{1}{b \ln s/\Lambda^2_{\overline{MS}}}, \tag{3.192}$$

where

$$b = \frac{1}{4}\left(11 - \frac{2}{3} N_f - \frac{1}{6} N_s\right),$$

and N_f and N_s are the numbers of quark and scalar triplets respectively.

In a rigorous sense, the value of the fundamental scale Λ in models with integral charged quarks differs from the relevant value in standard QCD owing to the additional scalar degrees of freedom. Choosing $\Lambda_{\overline{\text{MS}}} \simeq 100$ to 200 MeV as a reference point, which follows for example from deep inelastic lepton-nucleon scattering, we find that the contribution of the perturbational correction to the annihilation cross section within a scalar sector is significant at least in the energy range $\sqrt{s} \lesssim 10$ GeV.

This shows, perhaps, that there is a production threshold of the physical hadronic states that incorporate color scalar fields their

structure and it lies at higher energies. We shall demonstrate below that a theoretical estimate of the mass scale of hadrons (i.e. hadrons having one or more quarks substituted by color scalars) agrees with this conclusion.

3.7.3 Scalar Quarks and New Hadrons

The above analysis of spontaneously broken color symmetry was based on an introduction of fundamental color scalars. One major implication of such models is that a new hadron family made up from quarks and colored scalars strongly bound by QCD forces could exist. The experimental observation of states like this would be commensurable with the discovery of a new flavor, namely, scalar quarks.

As was shown above, all the available experimental data does not preclude the possibility that color symmetry could break down spontaneously, provided the scale of the breakdown in terms of mass is sufficiently small, e.g.

$$\langle\varphi\rangle = v \ll 1 \text{ GeV}.$$

Thus, scalar quarks in models with the spontaneously broken color symmetry should have small mass.

At first sight, low-mass colored scalars, or scalar quarks, should result in new hadrons that have masses of the order of the scale characteristic of low-energy hadron physics ($\sim$1 GeV). This, however, contradicts experiment.

Below we will show that thanks both to the large perturbational corrections in QCD with scalar quarks and to the nonperturbational effects due to a possible scalar condensate, i.e.

$$\langle\alpha_s^{1/b}\varphi^+\varphi\rangle \simeq -(1\text{ GeV})^2, \tag{3.193}$$

the mass of a new hadron would be of the order of several tens of a GeV. Thus, the new hadrons could, in principle, be discovered and investigated at accelerators like PETRA, LEP, SPS, and UNK in colliding beams experiments.

A hadron's properties are substantially governed by the structure of the vacuum state. Introducing low-mass scalar quarks into CQD changes the vacuum structure, since in addition to the well-known quark and gluon condensates [50] viz.

$$\langle\bar{q}q\rangle \simeq -(0.25\text{ GeV})^3, \quad q=u,\ d, \tag{3.194}$$
$$\langle\alpha_s G^a_{\mu\nu}G^a_{\mu\nu}\rangle \simeq +(0.5\text{ GeV})^4,$$

it emerges that there is a scalar condensate $\langle\varphi^+\varphi\rangle$.

The value of $\langle\varphi^+\varphi\rangle$ cannot be determined at the present time in a purely theoretical way. It is, like the quark and gluon field condensates, a phenomenological parameter. We do have, from dimen-

sional considerations, that $\langle \varphi^+ \varphi \rangle = c\Lambda^2$, where Λ is the fundamental scale (the inverse confinement radius) of QCD.

However, QCD coupling is known to be substantially different for different channels and is characterized (depending on the channels' quantum numbers) by different effective scales Λ (α).

The computation of the three-loop corrections to the total cross section of the annihilation of e^+e^--pairs into hadrons with fundamental scalar triplets shows that the effective scale in the production channel for the $\varphi^+\varphi^-$-pairs is about an order of magnitude larger than the corresponding scale in the quark-antiquark channel, i.e. $\Lambda(\varphi^+\varphi) \simeq$ $\simeq 9\Lambda(\bar{q}q)$ in the $\overline{\text{MS}}$ renormalization scheme. This plays a significant role in the subsequent discussion[10]. The mass of a bound state of a colored scalar is defined by the average $\langle \varphi^+ \varphi \rangle$ in the same way as the mass of a meson made up from quarks and antiquarks is defined mainly by the quark condensate.

Further we shall consider the bound states of the scalar and normal quarks to have the quantum numbers of the following local currents:

$$\begin{aligned}
&\rho_s = \varphi^+ \overleftrightarrow{\nabla}_\mu \varphi, \\
&\nabla_\mu = \partial_\mu - i\frac{g}{2}\lambda^a A_\mu^a - ig'\frac{2}{3}A_\mu^0, \\
&\pi_s = \alpha_s^{1/5}(\varphi^+\varphi), \\
&\chi_s = \alpha_s^{1/2b}(\varphi^+ q),
\end{aligned} \tag{3.195}$$

where $b = \frac{1}{4}\left(11 - \frac{2}{3}N_f - \frac{1}{6}N_s\right)$. The quantities $\alpha_s^{1/b}$, $\alpha_s^{1/2b}$ provide the renorm-invariance of the two-point Green functions that correspond to the currents π_s and χ_s in the main logarithmic approximation.

The current and a corresponding particle ρ_s which we call the vector phionium, have quantum numbers $J^{PC} = 1^{--}$. Like the ρ-meson, the ρ_s-particle may be produced in e^+e^- collisions and will manifest itself as a resonance.

The quantum numbers of the scalar phionium π_s are $J^{PC} = 0^{++}$. The bound state of normal and scalar quarks χ_s, has spin 1/2 and can be called the "white" quark. The electric charges of ρ_s and π_s are equal to zero and the charge on the χ_s depends on the flavor of the corresponding quark, for example, $Q_{\varphi^+u} = 0$, $Q_{\varphi^+d} = -1$, etc.

The properties of the bound states of the scalar quarks (3.195) can be studied using QCD's sum rules taking into account the nonperturbational effects.

---

[10] We should point out in this connection, what was mentioned above (see Sec. 3.3.2) when we discussed the annihilation model of meson decays concerning the dependence of the effective radius of the mesons' quark-antiquark systems on their quantum numbers [35, 36].

Using the well-known techniques of operator expansion and QCD's sum rules [49, 50] we present the basic relations:

$$\frac{1}{M^2}\int_0^\infty ds\, e^{-s/M^2} R_{\rho_s}(s) = \frac{1}{3}\Big\{1+\frac{4}{\pi}\alpha_s(M^2,\ \Lambda_1(\rho_s))$$

$$+\frac{32\pi^2}{M^2}\sigma\,[\alpha_s(M^2,\ \Lambda_2(\rho_s))]^{-1/b}+O(1/M^4)\Big\}, \tag{3.196a}$$

$$\frac{1}{M^2}\int_0^\infty ds\, e^{-s/M^2} R_{\pi_s}(s) = [\alpha_s(M^2,\ \Lambda_1(\pi_s))]^{2/b}$$

$$+\frac{32\pi^2}{M^2}\sigma\,[\alpha_s(M^2,\ \Lambda_2(\pi_s))]^{1/b}+O(1/M^4), \tag{3.196b}$$

$$\frac{1}{M^2}\int_0^\infty ds\, e^{-s/M^2} R_{\chi_s}(s) = [\alpha_s(M^2,\ \Lambda_1(\chi_s))]^{1/b}$$

$$+\left(1+\frac{1}{3\pi}\alpha_s\right)\frac{32\pi^2}{3M^2}\sigma+O(1/M^4), \tag{3.196c}$$

where

$$\sigma \equiv \langle \alpha_s^{1/b}\varphi^+\varphi\rangle, \qquad \alpha_s(M^2,\ \Lambda) \equiv \pi/b\ln\frac{M^2}{\Lambda^2}.$$

The quantities $R_{\rho_s}(s)$, $R_{\pi_s}(s)$, and $R_{\chi_s}(s)$ are positive definite spectral densities related to the two-point Green functions which have the corresponding currents ρ_s, π_s, and χ_s, e.g.

$$i\int \langle 0|\,T\,(\rho_s^\mu(x)\,\rho_s^\nu(0))\,|0\rangle\, e^{iqx}\, d^4x$$

$$=(q_\mu q_\nu - q^2 g_{\mu\nu})\,R_{\rho_s}(q^2), \tag{3.197}$$

etc.

The scale parameters Λ_i, which are computed allowing for the perturbational corrections, are equal to

$$\begin{aligned} &\Lambda_1(\rho_s)\cong g\gamma\Lambda_{\overline{MS}}, \quad \Lambda_2(\rho_s)\cong 2.7\,\gamma\Lambda_{\overline{MS}},\\ &\Lambda_2(\pi_s)\cong 3.2\gamma\Lambda_{\overline{MS}}, \quad \gamma\cong 1.33\end{aligned} \tag{3.198}$$

(γ being the Euler constant), respectively.

The sum rules for the new hadrons differ qualitatively from the similar quarkonium relations by the contribution of the non-perturbational effects arising here from the terms which are of the order of $1/M^2$, rather than of $1/M^4$. This fact has a simple explanation and is associated with the dimensions of the fields since dim $\langle\varphi^+\varphi\rangle = 2$, i.e. the term with the lowest dimension in the sum rules is proportional to $\langle\varphi^+\varphi\rangle/M^2$ for the scalar field. For the fermion fields, dim $\langle\bar{q}q\rangle = 3$ and the expansion starts from the $m_q\langle\bar{q}q\rangle/M^4$ and $\langle G_{\mu\nu}^2\rangle/M^4$

term. It is this difference which leads to the significantly larger mass of the new hadrons (the phioniums) as compared with the masses of quarkoniums.

An analysis of the sum rules (3.196) leads to the following estimates for the mass of a phionium:

$$m_{\rho_s'} \simeq (45 \pm 5)\,(-\sigma)^{1/2},$$
$$m_{\pi_s} - m_{\chi_s} \simeq \frac{1}{\sqrt{3}}\, m_{\rho_s}, \tag{3.199}$$

where the quantity σ varies within the limits

$$(0.5\ \ \mathrm{GeV})^2 < |\,\sigma\,| < (2\ \mathrm{GeV})^2.$$

Unfortunately, a numerical value for the scalar potential cannot be extracted from the experimental data available, since it does not significantly influence the parameters of the known resonances of quarks. Assuming, for example, that the value for the scalar condensate is related to the phionium scale ($\Lambda_1(\rho_s) \simeq 9\gamma\Lambda_{\overline{\mathrm{MS}}}$) in the same way as the quark condensate value is related to the quarkonium scale ($\Lambda_1(\rho) \simeq 1.4\gamma\Lambda_{\overline{\mathrm{MS}}}$), we can obtain the following for $\Lambda_{\overline{\mathrm{MS}}} \simeq 100\,\mathrm{MeV}$:

$$\langle\varphi^+\varphi\rangle^{1/2} = \langle\bar{q}q\rangle^{1/3}\cdot\frac{\Lambda_1(\rho_s)}{\Lambda_1(\rho)} \simeq 1.5\ \mathrm{GeV}. \tag{3.200}$$

Thus, the mass characteristic of the new hadrons is significantly larger than that of hadrons constructed from low-mass quarks and can be up to some tens of a GeV.

In a rigorous sense, a small value for the scalar condensate, say, $(-\sigma)^{1/2} \sim 0.5$ GeV, cannot presently be excluded by the theory. In this case the new particles would have been observed in the experiments in the accelerators now in operation. Note that the experimental search for long-lived ($\tau > 10^{-13}$ sec) particles in the e^+e^--annihilation processes [88] yields a limit for the "white" quark mass, namely,

$$m_{\chi_s} > 12\ \mathrm{GeV}.$$

In the model under consideration the lightest χ_s-particles must be absolutely stable by virtue of baryonic number conservation. When $SU^c(3) \times U(1)$-theory is embedded into grand unification models, the stability property of the lightest "white" quarks vanishes. In theories with "late" unification (such as $SU(5)$, $SO(10)$), the lifetime τ of χ_s-particles is about 10^{30} years, and in theories with small unification scales (for example, $SU^4(4)$, $SU^2(8)$, etc., [89, 90]) it is about 10^{-8}-10^{+6} sec.

As has been noted above, colored scalar particles with small Lagrangian masses occur naturally in theories which permit weakly broken color symmetry.

A question arises as to how color breakdown influences the properties of the new hadrons, if they do in fact exist.

In this case the vacuum average $\langle\varphi^{*a}\varphi_b\rangle$, in addition to a contribution from the color symmetry's scalar condensate, which has a negative sign (as was shown above), contains a positive addend due to the color breakdown, namely,

$$\begin{gathered}\langle\varphi^{*a}\varphi_b\rangle = \frac{1}{3}\,\delta_b^a\,\langle\varphi^{*\prime}\varphi^{\prime}\rangle + \frac{1}{2}\,v^a v_b,\\ \varphi_a = \varphi_a^{\prime} + \frac{1}{\sqrt{2}}\,v_a, \quad \langle\varphi_a^{\prime}\rangle = \frac{1}{\sqrt{2}}\,v_a \neq 0.\end{gathered} \tag{3.201}$$

Obviously, the properties of the new hadrons in theories with spontaneously broken and those with exact color symmetries are practically coincident when $v^2 \ll (\langle\varphi^{+}\varphi\rangle)$.

Yet the so-called U-boson, which occurs in the model and which together with the γ-quantum diagonalizes the mass matrix of the vector particles, is perhaps unobservable, i.e. is not a physical state when there is a weak color breakdown.

The experimental implications of theories with the spontaneously broken color symmetry are discussed in more detail in [53], to which we refer the reader.

References

1. M. Gell-Mann, *Phys. Lett.*, 8, 214 (1964).
2. G. Zweig, Preprint TH-104, CERN, Geneva (1964).
3. Y. Ne'eman, *Nucl. Phys.*, **26**, 222 (1961).
4. N. N. Bogolyubov, B. V. Struminskii, and A. N. Tavkhelidze, Preprint D-1968, JINR, Dubna (1965);
   A. N. Tavkhelidze, *High Energy Phys. Elem. Particles*, Vienna (1965), p. 753.
5. M. Y. Han and Y. Nambu, *Phys. Rev.*, **139B**, 1005 (1965).
6. P. G. Freund, *Nuovo Cim.*, 39, 765 (1965).
7. Y. Miyamoto, *Progr. Theor. Phys. Suppl. Extra*, No. 187 (1965).
8. M. Gell-Mann, P. Ramond, and R. Slansky, *Rev. Mod. Phys.*, **50**, No. 4, 721 (1978).
9. J. J. J. Kokkedee, *The Quark Model*, Benjamin, New York-Amsterdam (1969).
10. V. P. Shelest, G. M. Zinov'ev, and V. A. Miransky, *Models of strongly interacting elementary particles*, v. 1, Atomizdat, Moscow (1975) (in Russian).
11. N. N. Bogolyubov, in: *High energy physics. Theory of elementary particles symmetry*, Naukova Dumka, Kiev (1967), pp. 5-112 (in Russian).
12. B. V. Struminskii and A. N. Tavkhelidze, in: *High energy physics. Theory of elementary particles symmetry*, Naukova Dumka, Kiev (1967), pp. 625-638 (in Russian).
13. N. N. Bogolyubov *et al.*, Preprint P-2141, INR, Dubna (1965); see also in: *Problems of elementary particle physics*, Izd. Akad. Nauk SSSR, Erevan (1965), pp. 406-420 (in Russian).
14. P. N. Bogolyubov, *Ann. Inst. H. Poincaré*, v. VIII, 2 (1968).
15. A. Chodos *et al.*, *Phys. Rev.*, D**12**, 2060 (1975).

16. Y. Nambu, in: *Preludes in Theoretical Physics* (Ed. R. Gatto), New York, I. Wiley (1972).
17. H. Fritzsch and M. Gell-Mann, *Proc. XVI Int. Conf. High Ener. Phys.*, v. 2, NAL, Batavia (1972);
H. Fritzsch, M. Gell-Mann, and H. Leutwyller, *Phys. Lett.*, **74B**, 365 (1973).
18. C. N. Yang and R. L. Mills, *Phys. Rev.*, **96**, 191 (1954).
19. O. W. Greenberg, *Phys. Rev. Lett.*, **13**, 598 (1964).
20. O. W. Greenberg and A. M. L. Messiah, *Phys. Rev.*, **138B**, 1155 (1965).
21. A. B. Govorkov, in: *High Energy Physics. Theory of elementary particles symmetry*, Naukova Dumka, Kiev (1967), p. 770 (in Russian); *Ann. Phys. (N.Y)*, **53**, 349 (1969).
22. A. A. Logunov, M. A. Mestvirishvili, and V. A. Petrov, *Quantum Field Theory Fundamentals and Their Corollaries*, Nauka, Moscow (1977) (in Russian).
23. V. A. Matveev, R. M. Muradyan, and A. N. Tavkhelidze, Preprint P2-4578 JINR, Dubna (1969);
Elementary particles and atomic nucleus, v. 2 (1971), pp. 6-32;
Lett. Nuovo Cim., **5**, 907 (1972).
24. N. N. Bogolyubov, V. S. Vladimirov, and A. N. Tavkhelidze, *Theor. and Math. Phys.*, **12**, No. 1, 619 (1972);
Theor. and Math. Phys., **12**, No. 3, 839 (1972).
25. V. A. Matveev, R. M. Muradyan, and A. N. Tavkhelidze, *Lett. Nuovo Cim.*, **7**, 712 (1973).
26. S. Brodsky and G. Farrar, *Phys. Rev. Lett.*, **31**, 1153 (1973).
27. R. G. Arnold *et al.*, *Phys. Rev. Lett.*, **35**, 776 (1975).
28. V. A. Matveev, R. M. Muradyan, and A. N. Tavkhelidze, in: *IV Int. Seminar Probl. High Ener. Phys.* (1975) (ed. A. M. Baldin *et al.*), Preprint D1-9224, INR, Dubna (1975), p. 215;
S. Brodsky, in: *Few Body Probl. Nucl. Particle Phys.*, Les Presses del'Univ. Laval, Quebec (1975), p. 676.
29. V. A. Matveev, *Multiquark systems. Theory and experiment*, Lectures at High Energy Physics School, Gomel', BSSR, Preprint P2-12080, JINR, Dubna (1978), pp. 57-89.
30. J. C. Pati and A. Salam, *Phys. Rev.*, D8, 1240 (1973).
31. A. Yu. Ignat'ev *et al.*, *Theor. and Math. Phys.*, **47**, No. 2, 215 (1981).
32. N. N. Bogolyubov, V. A. Matveev, and A. N. Tavkhelidze, in: *Problems of Elementary Particle Theory*, Proc. Intern. Seminar, Varna, Preprint P2-4050, I JINR, Dubna (1968), p. 269.
33. A. Salam, *Int. Conf. High Ener. Phys.*, Tbilisi (1976).
34. V. A. Matveev, B. V. Struminskii, and V. G. Torgamadze, Preprint P-2821, JINR, Dubna (1965);
R. Gatto, L. Miani, and G. Preparata, *Phys. Rev. Lett.*, **16**, 377 (1966).
35. V. A. Matveev, B. V. Struminskii, and A. N. Tavkhelidze, Preprint P-2524, JINR, Dubna (1965).
36. R. Van Royen and V. F. Weisskopf, *Nuovo Cim.*, **50A**, 617 (1965); *Nuovo Cim.*, **51A**, 583 (1967).
37. V. A. Matveev, Preprint P2-3847, JINR, Dubna (1968).
38. A. N. Tavkhelidze, *Proc. XIV Conf. Particle Phys.* Univ. Brussels, Intanci (1968), pp. 145-154.
39. A. A. Logunov and A. N. Tavkhelidze, *Nuovo Cim.*, **29**, 380 (1963).
40. V. P. Shelest, Preprint 67-18, itp. ITP, Kiev (1967).
41. V. A. Matveev, R. M. Muradyan, and A. N. Tavkhelidze, Preprint E2-3498, JINR, Dubna (1967);
Preprint P2-3900, JINR, Dubna (1968).
42. R. N. Faustov, *Theor. and Math. Phys.*, **3**, No. 1, 152 (1970); *Ann. Phys.*, **78**, 176 (1973).
43. R. Dashen and M. Gell-Mann, *Phys. Rev. Lett.*, **17**, 340 (1966)

44. V. A. Matveev, *Quasipotential method in strong interaction theory at high energies*, Dr. Sc. Thesis, JINR, Dubna (1972) (in Russian).
45. V. R. Garsenashvili *et al.*, *Theor. and Math. Phys.*, **23**, No. 2, **310** (1975).
46. D. Gross and F. Wilczek, *Phys. Rev.*, *Lett.*, **30**, 1346 (1973).
47. C. Callan, R. Dashen, and F. Gross, *Phys. Rev.*, **D17**, 2717 (1978).
48. L. B. Okun, *Sov. Phys. Uspekhi*, **134**, 3 (1981).
49. A. A. Logunov, L. D. Solov'ev, and A. N. Tavkhelidze, *Phys. Lett.*, **24B**, 181 (1967);
N. V. Krasnikov, A. N. Tavkhelidze, and K. G. Chetyrkin, *Proc. Intern. Conf. Neutrino*-77, v. 2, Nauka, Moscow (1978), p. 189. *Theor. and Math. Phys.*, **35**, No. 2, 373 (1978).
50. L. B. Okun and M. A. Shifman, *Z. Phys.*, **8C**, 17 (1981);
M. A. Shifman, A. Vainshtein, and V. I. Zakharov, *Nucl. Phys.*, **B147**, 385, 448, **519** (1979);
M. Gell-Mann *et al.*, *Phys. Rev.* **175**, 2195 (1968);
M. Leutwyller, *Phys. Lett.*, **48B**, 45 (1979);
A. I. Vainshtein, V. I. Zakharov, and M. A. Shifman, *Sov. Phys. JETP Lett.*
51. L. B. Okun, M. B. Volloshin, and V. I. Zakharov, Preprint ITEP-79, ITEP, Moscow (1979).
52. A. De Rujula, R. C. Giles, and R. L. Jaffe, *Phys. Rev.*, **D17**, 285 (1978).
53. A. Yu. Ignant'ev *et al.*, *Theor. i Mat. Fiz.*, **52**, 142 (1982).
54. O. W. Greenberg, *Phys. Repts.*, **C32** (1978).
55. H. S. Green, *Phys. Rev.*, **90**, 170 (1953).
56. D. V. Volkov, *Sov. Phys. JETP*, **36**, 1560 (1959);
Sov. Phys. JETP, 38, **519 (1960)**.
57. O. LeinKlein, *J. Phys. Rad.*, **9**, 1 (1938).
58. N. N. Bogolyubov, A. A. Logunov, and I. T. Todorov, *Introduction to Axiomatic Quantum Field Theory*. Benjamin, London (1975), XXVI, 708 pp.
59. A. B. Govorkov, Preprint E2-3003, JINR, Dubna (1966).
60. A. B. Govorkov, *Teor. i Mat. Fiz.*, (1982) (in print).
61. A. A. Logunov, M. A. Mestverishvili, and Ngyen Van Hieu, *Phys. Lett.*, **25B**, 611 (1967).
62. R. P. Feynman, *Photon-hadron interactions*, Benjamin, Reading, Massachussetts (1972).
63. C. N. Yang, *High Energy Collisions*, Gordon and Breach, New York (1969), p. 509.
64. M. A. Markov, *Neutrino*, Nauka, Moscow (1964).
65. J. D. Bjorken, *Lecture in Varenna School*, Course 41, Varenna (1967).
66. L. I. Sedov, *Similarity and dimensional techniques in mechanics*, Nauka, Moscow (1972).
67. V. A. Matveev, R. M. Muradyan, and A. N. Tavkhelidze, Preprints P2-4578, P2-4543, P2-4824, P2-, JINR, Dubna (1969).
68. J. Christensen *et al.*, *Intern. Symp. Electron Photon Inter. High. Ener.*, Daresbury (1969), p. 382.
69. S. M. Berman, J. D. Bjorken, and J. B. Kogut, *Phys. Rev.* D4, 3388 (1971).
70. D. Sivers, S. Brodsky, and Blankenbeckler, *Phys. Repts.*, **23C** (1976).
71. A. N. Tavkhelidze, *XVIII Int. Conf. High Ener. Phys.*, Tbilisi (1976).
72. S. V. Goloskokov and V. A. Matveev, *Lectures at XIV Intern. School High Ener. Phys.*, Preprint D2-81-158, JINR, Dubna (1981), pp. 205-226.
73. A. N. Efremov and A. V. Radyushkin, Preprint E2-11983, JINR, Dubna (1978).
74. S. Brodsky and G. P. Lepage, *Phys. Rev. Lett.*, **43**, 545 (1979).
75. V. L. Chernyak, A. P. Zhitnitskii, and V. G. Serbo, *Pis'ma Zh. Teor. i Eksp. Fiz.* 26, 760 (1977).
76. V. A. Matveev, P. M. Muradyan, and A. N. Tavkhelidze, *Theor. and Math. Phys.*, **40**, No. 3, 329 (1979).
77. V. A. Matveev and I. K. Solov'ev, Preprint P-80-742, JINR, Dubna (1980).
78. S. Mandelstam, *Phys. Rev.*, **175**, 1580 (1968).

79. I. Ya. Aref'eva and A. A. Slavnov, *Lectures at XIV Intern. School Phys. High Ener.*, Preprint D2-81-1581, JINR, Dubna (1981), pp. 36-101.
80. V. A. Matveev, R. M. Muradyan, and A. N. Tavkhelidze, Preprint E2-8048, JINR, Dubna (1974).
81. S. V. Goloskokov *et al.*, Preprint P2-81-420, JINR, Dubna (1981).
82. V. A. Matveev, L. A. Slepchenko, and A. N. Tavkhelidze, *Phys. Lett.*, **100B**, 75 (1981).
83. I. S. Avaliani, V. A. Matveev, and L. A. Slepchenko, Preprint P2-81-401, JINR, Dubna (1981).
84. A. M. Baldin, in: *High Energy Physics and Nuclear Structure* (ed. by D. E. Nagy *et al.*). Amsterdam American Institute of Physics, New York (1975), p. 621.
85. S. V. Goloskokov *et al.*, Preprint P2-10142, JINR, Dubna (1976).
86. V. A. Matveev and P. Sorba, *Nuovo Cim. Lett.*, **20**, 443 (1977).
87. S. G. Goriskii *et al.*, *Proc. IV Intern. Conf. Quantum Field Theor. High Ener. Phys.*, IHEP, Protvino (1981).
88. W. Bartel *et al.*, *Z. Phys.*, **C6**, 295 (1980).
89. J. C. Pati and A. Salam, *Phys. Lett.*, **58B**, 333 (1975).
90. Yu. F. Pirogov, Preprint 80-78, IHEP, Protvino (1980).

AUTHOR INDEX

SUBJECT INDEX

CONTENTS